مدخل حديث للإحصاء والاحتمالات

إعداد
دكتورة / ثروت محمد عبد المنعم



مدخل حديث للإحصاء والاحتمالات

إعداد

د.ثروت محمد عبد المنعم

أستاذ مساعد بكلية العلوم بالدمام

قسم الرياضيات

المملكة العربية السعودية

الطبعة الأولى ١٤٢١ هـ ـ ٢٠٠٠ م



مكتبة الأنجلو المصرية

١٦٥ شارع محمد فريد - القاهرة

بسم الله الرحمن الرحيم

وما أوتيتم من العلم إلا قليلا

(صدق الله العظيم)

الإهداء

إلى أمي التي علمتني العطاء.

إلى والدي الذي علمني الخلق الكريم.

إلى أستاذي ا.د.أحمد حسن الموازيني الذي علمني أن الثقة بالنفس أول خطوات النجاح.

إلى أستاذي ا.د.سمير كامل عاشور و أستاذتي ا.د.الهام شكري اللذان علماني أن تشجيع الأستاذ لتلميذه دافع قوي له على التقدم.

إلى أستاذي المرحوم ا.د.علي يحي الذي علمني أنه لا يوجد صعب في العلم. إلى أخوتي فهم سندى في الحياة

إلى أخواتي في الله د.سارة و د.أميرة اللتان علمتاني أن الأخوة ليست فقط في الرحم.

إلى كل من شجعني في حياتي و أعطاني دفعة نحو الأمام.

بسم الله الرحمن الرحيم

تقديم

الحمد لله رب العالمين والصلاة والسلام على أشرف المرسلين محمد وعلى صحبه أجمعين وبعد..

بعون الله سبحانه وتعالى نقدم هذا الكتاب للمكتبة العربية وللباحثين العرب في المجالات المختلفة تلبية لنداء التعريب الذي يتبناه الكثير من العلماء و المثقفين. يهدف هذا الكتاب إلى تقديم الطرق الإحصائية للدارسين والطلاب المتخصصين في الإحصاء مع عدم الخوض في الكثير من الإشتقاقات الرياضية وكذلك البعد عن التعمق في النظريات الإحصائية والتركيز على الفهم والتطبيق. وقد استندت في الشرح النظري على الكثير من الأمثلة التطبيقية المدعمة بالتمارين التي تلي كل فصل في الكتاب لتؤكد الفهم واختبار القدرة على التطبيق.

معظم الكتب الإحصائية تبدأ بعرض الطرق الوصفية ثم بموضوع الاحتمالات كجزء منفصل ثم بطرق الاستدلال الإحصائي. يقدم هذا الكتاب المحتويات بأسلوب مختلف حيث يبدأ بموضوع الاحتمالات ثم يربطه مباشرة بالطرق الإحصائية سواء في مجال الإحصاء الوصفي أو الإستدلالى. هذا وقد راعيت أن تكون المعادلات بالرموز اللاتينية وذلك للتيسير على القارئ في الاستفادة من المراجع الأجنبية. كما استعنت في إعداد هذا الكتاب أولا بخبرتي الطويلة في تدريس الإحصاء وفى الأبحاث والاستشارات الإحصائية وثانيا بكثير من المراجع العربية والأجنبية والمعطاة أسمائهم في نهاية الكتاب. هذا وقد تم أعداد الرسوم البيانية باستخدام البرامج الجاهزة التالية : **SPSS , Harvard , Mathematica , Statistica** بالإضافة إلى برنامج الراسم التابع للنوافذ 97.

يصلح هذا الكتاب كمقررين لطلبة الجامعات في شتى التخصصات حيث يدرس الفصول الخمسة الأولى كمقرر وباقي الفصول كمقرر ثاني، كما يصلح لأن يكون مرجعا لأي باحث. هذا ويمكن للباحثين المستفيدين من هذا الكتاب والذين يعملون في المجالات التطبيقية تخطي براهين النظريات الموجودة في الكتاب.

وأخيرا أود أن اعبر عن شكري الجزيل لكل الذين شجعوني وساعدوني بصورة أو بأخرى على إخراج هذا الكتاب وأخص بالذكر زميلاتي وأخوتي في الله د. سارة السدحان ود. أميرة الضاوي اللتين قدمتا لي الكثير من التشجيع والعون سواء في الرسوم البيانية أو في المراجعة اللغوية وأدعو لهما وللجميع بالتوفيق.

كما اتوجه بالشكر لمكتبة الانجلو المصرية التي أتاحت لى الفرصة لنشر هذا العمل العلمي والله ولى التوفيق، وإني أرحب بكل نقد بناء يهدف إلى الأفضل، فالكمال لله وحده.

تـقـديـم

الكتاب الذي بين يدي القارئ عمل مميز يتناول أحد فروع الرياضيات التطبيقية التي تختص بتقديم الطرق و الأساليب التي تستخدم في تحليل الظواهر .

استمعت بقراءة هذا الكتاب العلمي الذي صيغت الموضوعات الواردة فيه بطريقة مشوقة سلسة تجذب القارئ إلى أسلوبه العلمي المفصل ، وهو متنوع الأمثلة مدعم بتمارين شاملة في نهاية كل فصل و مستنداً لعدد كبير من المراجع القيمة ليوسع من مدارك القارئ العربي ولذلك فهو لا غنى عنه لطلاب كليات العلوم بمختلف تخصصاتها كما يمكن أن يخدم طلاب الدراسات العليا للتخصصات الحيوية.

و عبر هذه الصفحات يتضح المجهود الكبير الذي بذلته الدكتورة ثروت محمد عبد المنعم ويؤهلها في ذلك ما نالته من خبره المشاركة الفعالة في البحث و التدريس عبر خمس عشر عاما بحمد الله و توفيقه و التي أتمنى لها مستقبلاً زاهراً.

أ. د الهام شكري

وكيلة معهد الإحصاء - جامعة القاهرة

المحتويات

الفصل العاشر : الانحدار و الارتباط

الفصل الحادي عشر : تحليل التباين

الفصل الثاني عشر : الاختبارات اللامعلمية

الفصل الأول

مقدمه

Introduction

يهتم الإحصائي ، أساسا ، بالنواتج التي يحصل عليها الباحثون من إجراء الأبحاث العلمية. فعلى سبيل المثال قد يهتم بعدد الحوادث التي تقع في تقاطع ما، عدد الأفراد في الأسرة، عدد البكتريا في لتر من الماء النقي، عدد المصابين بالسرطان، كمية الحديد التي تستهلكها الإناث البالغات، عدد المصابين بتسوس للأسنان، أطوال مجموعة من الأشخاص...الخ ، أي أنه يهتم بقراءات (مشاهدات) قابلة للعد أو للقياس. سوف نشير إلى المعلومات **information** المسجلة في صورتها الأصلية التي جمعت بها بالبيانات الخام **raw data** . يهتم الإحصائي بالطرق المختلفة لوصف كمية كبيرة من البيانات الخام وذلك بغرض الوصول إلى معلومات لم تكن معلومة له. أو قد يهتم بالوصول إلى استنتاجات أو اتخاذ قرارات عن فئة كبيرة من البيانات وذلك بالاعتماد على فئة جزئية منها.

يهتم علم الإحصاء بالطرق المستخدمة في جمع و عرض و وصف و تحليل و تفسير البيانات. كل المشاكل التي تستخدم الطرق الإحصائية يمكن أن تنتمي إلى مجال الإحصاء الوصفي **descriptive statistics** أو الإحصاء الاستقرائي **inductive statistics** ، أي معالجة للبيانات تؤدى إلى تنبؤات أو استدلالات تخص مجموعة كبيرة من البيانات تجعلنا في مجال الإحصاء الإستقرائي. ومن ناحية أخرى إذا كان اهتمامنا منحصر في البيانات التي في حوزتنا ولا توجد أي محاولات لتعميمها إلى فئة أكبر من البيانات، فإننا نكون في مجال الإحصاء الوصفي. ويمكن توضيح الفرق بين الإحصاء الوصفي والإحصاء الاستقرائي من المثال التالي : فإذا كان اهتمامنا بفئة من القيم الممثلة لكمية الصادرات لسلعة ما في منطقة ما خلال العشرين سنة الماضية، فإن أي قيمة تصف البيانات، مثل متوسط كمية الصادرات في العشرين سنة الماضية ، هي قيمة في مجال الإحصاء الوصفي. في هذه الحالة لم نحاول أن نذكر أي شيء عن كمية الصادرات لأي سنة أخرى غير السنوات العشرين والتي حصلنا منها على المعلومات. فإذا كان متوسط كمية الصادرات في هذه المنطقة في العشرين سنة الماضية **1.34** طن واستطعنا أن نضع جملة تنص على أن كمية الصادرات في السنة القادمة سوف تتراوح بين **1.2** إلى **1.7** طن، في هذه الحالة نكون قد عممنا ووضعنا أنفسنا في مجال الإحصاء الإستقرائي.

التعميم في الإحصاء الاستقرائي يخضع للأحداث غير المؤكدة لأننا نهتم فقط بمعلومات جزئية تم الحصول عليها من فئة جزئية من البيانات موضع الدراسة. للتعامل مع الأحداث غير المؤكدة فإن فهمنا لنظرية الاحتمالات يكون ضروريا. في هذا الكتاب سوف نقدم بعض المفاهيم الأساسية في نظرية الاحتمالات وذلك في الفصل الثالث والرابع. وبما أنه يمكن عرض الاحتمالات بصورة أفضل وباستخدام صيغ الفئات لذلك سوف نقدم بعض الأساسيات عن

الفئات وخوصها في الفصل الثاني . أما بالنسبة للرياضيات المطلوبة لمقرر الإحصاء فـلا تتعـدى بعض الصيغ الأساسية في مقرر الجبر الذي يدرس في الجامعات . بعض الصيغ الرياضية المفيـدة في مجال الإحصاء سوف نتناولها في الفصل الثاني .

يتناول هذا الكتاب الأسس العامة لمبادئ الإحصاء والإحتمالات . تتنـاول الفصـول الأولى من الكتاب بعض الصيغ الضرورية لدراسة الإحصاء والاحتمالات، والمتغيرات العشوائية، وبعض التوزيعات الاحتمالية، وتطبيقاتها، وطرق عرض ووصف البيانات وما يتعلـق بهـا مـن مقاييس الالتواء والتفلطح . ويتطرق الكتاب في الفصول الأخيرة إلى بعض توزيعـات المعاينـة واستخدامها في إيجاد فترات الثقة واختبارات الفروض وتحليل التباين والارتباط والإنحدار . أمـا الفصل الأخير فيهتم بالاختبارات اللامعلمية .

يقدم هذا الكتاب المعلومات بالوسيلة التي تفيد كل التخصصـات سـواء في الطـب، الزراعة، الأعمال . . . الخ . فالطرق الأساسية لجمع وعرض وتحليل البيانات واحدة بصـرف النظر عن مجال التطبيق . فعلى سبيل المثال قد يقوم الكيميائي بإجراء تجربة ما مستخدما أربع طرق بغرض قياس كمية منتج ما، موضع الدراسة، ثم تحليل النتائج باستخدام الطرق الإحصائية . هذه الطرق نفسها، يمكن استخدامها لتحليل البيانات التي تمثل عدد الوحدات المنتجة التالفة باستخدام أربع آلات مختلفة أو لتحليل البيانات التي يتم الحصول عليها من قياس كمية المحصول عند اختبار أربعة أسمدة مختلفة . كثير من الطرق المصممة للتطبيقات الزراعية أثبتـت كفاءتهـا في مجـالات أخرى .

في الحقيقة يعتبر الإحصاء أداة قوية جدا وضرورية إذا ما استُخدمت بطريقة صحيحـة . ولذلك لابد من تطبيق الطرق الصحيحة والأكثر كفاءة للظروف المعطاة وذلك من أجل الحصول على أقصى معلومات من البيانات المتوفرة . الطرق المستخدمة لتحليل فئة من البيانات تعتمـد بدرجة كبيرة على الطريقة المستخدمة في جمع البيانات . ولهذا السبب يكون من المفضل استشارة الإحصائي من بداية التخطيط للبحث، حتى الوصول إلى النتائج وتحليلها وتفسيرها .

الفصل الثاني

بعض المفاهيم الرياضية

Some Mathematical Concepts

يتناول هذا الفصل بعض المفاهيم الأساسية في الرياضيات والضرورية لدراسة الإحصاء .

١-١) الفئات Sets

الفئة هي مجموعة من الأشياء المعرفة تماما . على سبيل المثال الأنهار في أفريقيا، الأعــداد الصحيحة الموجبة ...,1,2,3 ، الحروف الأبجدية من أ إلى ى كلها تمثل فئات . الأشياء التي تتكون تسمى عناصر elements أو أعضاء members في الفئة . عادة يمكــن تمثيــل ف الإنجليزية الكبيرة مثل A , B , C , Y . هناك طريقتان لوصف الفئة، إذا احتوت الفئة على عدد محدود من العناصر بحيث يمكن عمل قائمة بهذه العناصر، فعلى سبيل المثال الفئــة A والتي تتكون من العناصر 2,5,6,7 يمكن كتابتــها علــى الشــكل $A = \{2,5,6,7\}$ أو الفئة B ، والتي تمثل نواتج إلقاء زهرة نرد يمكن كتابتها على الشــكل $B = \{1,2,3,4,5,6\}$. يمكن وصف الفئة بجملة statement أو قاعدة rule فعلى سبيل المثال إذا كانت Y تمثل الفئة من كل الأشخاص في العالم وإذا كان y عنصر اختياري في Y ، فإنه يمكن كتابة الفئة Y علــى الشكل :

$$Y = \{y \mid y \text{ is a person living in the earth}\}.$$

ويعتمد وصف الفئة، سواء بقائمة أو قاعدة على نوع المشكلة موضع الدراسة . على سبيل المثال يكون من الصعوبة وضع قائمة بعناصر الفئة من الأزهار الحمراء في العالم . ومن ناحية أخــرى لا توجد قاعدة بسيطة لوصف الفئة :

$$D = \{\text{family, book, flower}\}.$$

باستخدام صيغ الفئة فإن الرمز $\in$ يعنى " أنه عضو في " أو " ينتمي إلى " والرمز $\notin$ يعنى " ليس عضو في " أو " لا ينتمي إلى " . إذا كان x عنصر في الفئة A و y ليس عنصــر في A فإن $x \in A , y \notin A$.

مثال (٢-١) ليكن $A = \{2,4,6,8\}, \; B = \{x|x \text{ is an integer divisible by 7}\}$ فإن $8 \in A \;,\; 3 \notin A \;,\; 49 \in B$

تعريف : تتساوى فئتين إذا احتوت الاثنتان بالضبط على نفس العناصر .

إذا كانت الفئة A تساوى أو تماثل الفئة B ، فإن كل عنصر ينتمي إلى A ينتمــي إلى B ، وكل عنصر ينتمي إلى B ينتمي إلى A . وسوف نرمز لهذا التســاوي بكتابــة $A = B$. في بعض الأحيان، إذا كانت إحدى الفئتين A أو B تحتوى على عنصر واحد على الأقل لا ينتمي إلى الاثنين، فإننا نقول أن الفئتين غير متساويتين، وفي هذه الحالة نكتب $A \neq B$.

مثال (٢-٢) لتكن $A = \{1,4,5\}, B = \{4,6,9\}, C = \{1,5,4\}$ فإن :-

$$A = C \ , B \neq C.$$

ويجب أن نتذكر أن الفئات لا تتغير عندما نغير ترتيب العناصر ۰

تعريف : الفئة الخالية **empty** أو فئة العدم **null set** هي الفئة التي لا تحتوى على أى عناصر ويرمز لها بالرمز ϕ ۰

إذا كانت :

$$A = \{x| x \text{ is a letter before A in the alphabet}\}$$

و

$$B = \{x|x^2 = 4, x \text{ is an odd number}\}$$

فإن B و A فئتان خاليتان ۰

لتكن الفئة $A = \{1,4\} , B = \{1,4,7,8,9\}$ ۰يلاحظ أن كل عنصر في الفئة A هو أيضا عنصر في الفئة B ۰ الفئة A تسمى فئة جزئية **subset** من B وتكتب $A \subset B$ ۰

تعريف : إذا كان كل عنصر في الفئة A هو عنصر في الفئة B فإن A تسمى فئة جزئية من B ۰

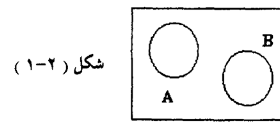
بناء على ذلك، فإن أى فئة تعتبر فئة جزئية من نفسها ۰

في كثير من المناقشات كل الفئات تعتبر فئات جزئية من فئة واحدة خاصة، هذه الفئة تسمى الفئة الشاملة ويرمز لها بالرمز U ۰

مثال (٢-٣) كل الفئات الجزئية من الفئة الشاملة $U = \{4,5,6\}$ هي :- $\{4\},\{5\},\{6\},\{4,5\},\{4,6\},\{5,6\},\{4,5,6\},\{\phi\}$ أى أن عدد الفئات الجزئية من الفئة الشاملة التي عدد عناصرها n هو 2^n من الفئات الجزئية ۰

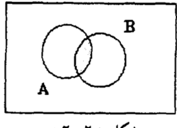
(٢-٢) عمليات الفئات Sets Operations

سوف نستعرض بعض العمليات على الفئات والتي تنتج فئات جديدة ۰ هذه الفئات الجديدة تعتبر فئات جزئية من الفئة الشاملة ۰ العلاقة بين الفئات الجزئية والفئة الشاملة يمكن توضيحها بما يسمى شكل فن **Venn diagram** ۰ تمثل الفئة الشاملة بمستطيل و الفئات الجزئية بدوائر داخل المستطيل كما في شكل (٢-١) ۰



شكل (٢-١)

سوف نرمز لتقاطع الفئتين B , A بالرمز $A \cap B$ • العناصر في $A \cap B$ لابد أن ينتمـــي إلى كل من B , A • يوضح الشكل (٢-٢) الفئة $A \cap B$ بالجزء المظلل.



شكل (٢-٢)

مثال (٢-٤) إذا كان $A = \{1,7,8,9,10\}, B = \{2,8,9,10,11\}$

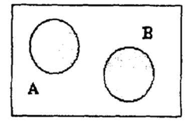
فإن $A \cap B = \{8,9,10\}$

مثال (٢-٥) إذا كان :-

$$A = \{x | x \text{ is an integer and } 1 \le x \le 6\},$$
$$B = \{y | y \text{ is an integer greater than } 4\}$$

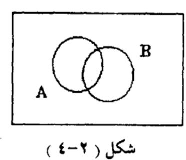
فإن $B = \{5,6,7,8,...\}, A \cap B = \{5,6\}$

تعريف : إذا كان $A \cap B = \phi$، يقال للفئتين B , A أنهمـــا منفصلتـــان **disjoint**، أي لا يوجد أي عناصر مشتركة بينهما كما في شكل (٢-٣) •



شكل (٢-٣)

تعريف : الاتحاد بين فئتين B , A هو الفئة من العناصر التي تنتمي إلى A أو B أو كلاهما • سوف نرمز للاتحاد بين A و B بالرمز $A \cup B$ • يوضح الشـــكل (٢-٤) الفئـــة $A \cup B$ بالجزء المظلل.

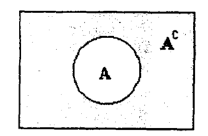


شكل (٢-٤)

مثال (٢-٦) إذا كان :-

$A \cup B = \{1,5,7,8,9,10,12\}$ فإن $A = \{1,5,7,8\}$, $B = \{8,9,10,12\}$

إذا كانت A فئة جزئية من الفئة الشاملة U ، فإن مكمل الفئة A هو الفئة من العناصر في U والغير موجودة في A . سوف نرمز لمكمل الفئة A بالرمز A^c . ويوضح شكل (٢-٥) الفئة المكملة بالجزء المظلل .



شكل (٢-٥)

مثال (٢-٧) بفرض أن :-

$A^c = \{8,9,10,12\}$ فإن $U = \{1,5,7,8,9,10,12\}$, $A = \{1,5,7\}$

هناك عدة نتائج من التعاريف السابقة مثل :-

$$A \cap \phi = \phi$$
$$A \cup \phi = A$$
$$A \cap A^c = \phi$$
$$A \cup A^c = U$$
$$U^c = \phi$$
$$\phi^c = U$$
$$(A^c)^c = A$$

(٢-٣) صيغ الجمع **Summation Notation**

سوف نحتاج في التحليل الإحصائي للبيانات إلى جمع مجموعة من الأعداد . إذا كانت x_i تمثل أى قيمة من القيم التالية التي عددها n : $x_1, x_2, \dots, x_n$ والتابعة للمتغير X . الحرف i في x_i ، الذي يأخذ أي من الأرقام $1,2,\dots,n$، يسمى الدليل **subscript** أو **index** . من الواضح أنه يمكن استخدام أي حرف غير i مثل s, p, q, j, k . الصيغة

$\sum_{i=1}^{n} x_i$ تستخدم لتمثيل جمع كل قيم x_i من $i = 1$ إلى $i = n$ • الرمز $\sum$ يسمى سيجما • وعلى ذلك فإن :

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n .$$

للتبسيط عادة يكتب المجموع بأشكال كثيرة مثل $\sum x$ أو $\sum x_i$ أو $\sum_{i} x_i$ •إذا كان :-

$x_1 = 2 \ , \ x_2 = 3 \ , \ x_3 = 5 \ , \ x_4 = 1 \ , \ x_5 = 2$ فإن :-

$$\sum_{i=1}^{5} x_i^2 = (2)^2 + (3)^2 + (5)^2 + (1)^2 + (2)^2 = 43.$$

مثال (٢-٨) إذا كان $x_1 = 2 \ , \ x_2 = 5 \ , \ x_3 = 6 \ , \ x_4 = 3 \ , \ x_5 = 4$ أوجد :-

(ا) $\sum_{i=2}^{5} (x_i - 1)$ (ب) $\sum_{i=3}^{5} 4x_i$

الحل •

(ا)
$$\sum_{i=2}^{5} (x_i - 1) = (5-1) + (6-1) + (3-1) + (4-1)$$
$$= 4 + 5 + 2 + 3 = 14.$$

(ب)
$$\sum_{i=3}^{5} 4x_i = (4)(6) + (4)(3) + (4)(4) = 52.$$

مثال (٢-٩) إذا كان $x_1 = 5 \ , \ x_2 = 4 \ , \ x_3 = 3 \ , \ x_4 = 1$ أوجد :-

(ا) $\sum_{i=1}^{4} (x_i - c)$ (ب) $\sum_{i=1}^{3} (c - i + 2)$ (جـ) $\sum_{i=1}^{3} x_i - 1$ حيث c عدد حقيقى•

الحل •

(ا)

$$\sum_{i=1}^{4} (x_i - c) = (x_1 - c) + (x_2 - c) + (x_3 - c) + (x_4 - c)$$
$$= (5 - c) + (4 - c) + (3 - c) + (1 - c) = 13 - 4c.$$

(ب)

$$\sum_{i=1}^{3} (c - i + 2) = (c - 1 + 2) + (c - 2 + 2) + (c - 3 + 2)$$
$$= 3c.$$

(جـ)

$$\sum_{i=1}^{3} x_i - 1 = (5+4+3) - 1 = 11.$$

(٢-٤) بعض النظريات المفيدة للجمع Useful Theorems Relating to Sums

نظرية (٢-١) إذا كان c عدد حقيقي فإن :-

$$\sum_{i=1}^{n} c = nc.$$

البرهان :-

$$\sum_{i=1}^{n} c = c + c + \ldots + c.$$

حيث أن c تتكرر n من المرات فإن :-

$$\sum_{i=1}^{n} c = nc.$$

مثال (٢-١٠)

$$\sum_{i=1}^{5} 4c = 5(4c) = 20c.$$

مثال (٢-١١)

$$\sum_{i=1}^{4} (5c - 3) = 4(5c - 3).$$

نظرية (٢-٢) إذا كان c عدد حقيقي فإن :-

$$\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i.$$

البرهان :-

$$\sum_{i=1}^{n} cx_i = cx_1 + cx_2 + \ldots + cx_n$$
$$= c(x_1 + x_2 + \ldots + x_n)$$
$$= c\sum_{i=1}^{n} x_i.$$

مثال (٢- ١٢) للبيانات في مثال (٢-٩) :-

$$\sum_{i=2}^{4} 7x_i = 7(4+3+1) = 56.$$

نظرية (٢-٣)

$$\sum_{i=1}^{n}(x_i + y_i + z_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} z_i.$$

البرهان :-

$$\sum_{i=1}^{n}(x_i + y_i + z_i) = x_1 + y_1 + z_1 + x_2 + y_2 + z_2 + x_3 + y_3 + z_3$$

$$+ \ldots + x_n + y_n + z_n$$

وبإعادة التجميع :-

$$\sum_{i=1}^{n}(x_i + y_i + z_i) = (x_1 + x_2 + \ldots + x_n) + (y_1 + y_2 + \ldots + y_n)$$

$$+ (z_1 + z_2 + \ldots + z_n) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} z_i.$$

مثال (٢-١٣) للبيانات في مثال (٢-٩) :-

$$\sum_{i=1}^{4}(x_i^2 + cx_i + 6) = \sum_{i=1}^{4} x_i^2 + \sum_{i=1}^{4} cx_i + \sum_{i=1}^{4} 6$$

$$= \sum_{i=1}^{4} x_i^2 + c\sum_{i=1}^{4} x_i + 4(6)$$

$$= (5)^2 + (4)^2 + (3)^2 + (1)^2 + c(5+4+3+1) + 24$$

$$= 51 + 13c + 24 = 75 + 13c.$$

مثال (٢-١٤)

$$\sum_{i=1}^{3}(c^2 + 7i) = \sum_{i=1}^{3} c^2 + \sum_{i=1}^{3} 7i$$

$$= 3c^2 + 7\sum_{i=1}^{3} i$$

$$= 3c^2 + 7(1+2+3)$$

$$= 3c^2 + 42.$$

٢-٥) الدالة Function

تعريف : بفرض أن X , Y فئتان غير خاليتين القاعدة التي تعين لكل عنصر $x \in X$ عنصر وحيد $y \in Y$ تسمى دالة من X إلى Y .

العنصر الوحيد y المعين بدالة ما والمناظر لعنصر محدد x يقال له قيمة الدالة عند x ويرمز للدالة بالرمز h, p, y, f ... الخ.

مثال (٢-١٥) إذا كانت X , Y فئتين من الأعداد حقيقية وكان $y = x + 3$ فإن y دالة في x . عندما $x = 1$ فإن $y = 4$ وعندما $x = -3$ فإن $y = 0$ وهكذا.

بفرض أن f دالة ما معطاة فإن الفئة X التي تعين الدالة f لكل عنصر من عناصرها عنصرا وحيدا $y \in Y$ يقال لها مجال domain الدالة . الفئة التي عناصرها العناصر المناظرة $y \in Y$ المعينة بالدالة f يقال لها مدى range الدالة f .

إذا كانت الدالة f تعين قيمة y في مداها لعنصر x في مجالها فإننا نكتب $y = f(x)$ وتسمى y قيمة f عند x . بفرض أن $y = 5x + 3$ وباستخدام صيغة الدالة فإننا يمكن أن نكتب $f(x) = 5x + 3$ وعلى ذلك $f(6) = 5(6) + 3 = 33$.

مثال (٢-١٦) إذا كانت $g(x) = x^2$ المطلوب إيجاد (ا) $g(2)$ (ب) المجال والمدى للدالة $g(x)$.

الحل.

(ا) $g(2) = 2^2 = 4$ (ب) المجال للدالة $g(x)$ هو جميع الأعداد الحقيقية والمدى جميع الأعداد الحقيقية الغير سالبة.

مثال (٢-١٧) إذا كانت $p(x) = \frac{1}{x^2}$ المطلوب إيجاد (ا) $p(3)$ (ب) المجال والمدى للدالة $p(x)$.

الحل. (ا) $p(3) = \frac{1}{3^2} = \frac{1}{9}$ (ب) المجال للدالة $p(x)$ هو جميع الأعداد الحقيقية ما عدا $x = 0$ والمدى جميع الأعداد الحقيقية الغير سالبة ما عدا $y = 0$.

تمارين

– ١ – أكتب عناصر الفئات التالية :–

(ا) الفئة من الأعداد الصحيحة بين 9 , 13 .

(ب) الفئة من الأعداد الصحيحة أقل من 9 .

(جـ) الفئة من الأعداد الصحيحة بين 1 و 20 القابلة للقسمة على 3 .

- ٢ - أكتب عناصر كل من الفئات التالية :-

(ا) الفئة $A = \{x | x^2 + x - 6 = 0\}$

(ب) فئة نواتج إلقاء زهرة نرد وعملة في نفس الوقت .

(جـ) الفئة $B = \{x | 2x - 4 = 0 \text{ and } x > 5\}$

- ٣ - بفرض أن $A = \{x | 3x = 6\}$ هل الجملة $A = 2$ صحيحة ؟

- ٤ - لتكن $E = \{3 , 5 , 6 , 8 , 9\}$ أى من الجمل التالية صحيح وأي خطأ ؟

(ا) $5 \in E$ (ب) $7 \in E$ (جـ) $9 \notin E$ (د) $\phi \subset E$

- ٥ - أي الفئات التالية تتساوى ؟

$\{t, s, r\}, \{r, s, t\}, \{t, r, s\}, \{s, r, t\}$.

- ٦ - أكتب عناصر كل من الفئات التالية :

(ا) الفئة $A = \{x | x \text{ is an integer and } 9 \le x \le 13\}$

(ب) الفئة $C = \{x | x + 2 = 0\}$

(جـ) الفئة $D = \{y | y^2 + 3y = 28\}$

- ٧ - لتكن $U = \{x, y, 6, z\}$ أكتب كل الفئات الجزئية من U .

- ٨ - أكتب كل الفئات الجزئية من $U = \{\text{man , woman, baby, home}\}$

- ٩ - لتكن $U = \{1,2,\ldots,9\}, A = \{1,2,3,4\}, B = \{2,4,6,8\}, C = \{3,4,5,6\}$

أوجد : - (ا) A^c (ب) $A \cap B$ (جـ) $(A \cap C)^c$ (د) $A \cup B$

- ١٠ - لتكن $U = \{10,11,12,13,14,15\}, A = \{10,11,12,13\}, B = \{12,13,14\}$

أوجد :- (ا) $A \cup B$ (ب) $A \cap B$ (جـ) A^c

- ١١ - إذا كانت :-

$$A = \{x | x \text{ is an integer and } x > 17\},$$
$$B = \{x | x \text{ is an integer and } 12 < x < 20\}$$

أوجد $A \cap B$

-١٢ - إذا كانت :-

$$A = \{x | x^2 - 7x = 8\},$$
$$B = \{x | 2x - 16 = 0\}$$

أوجد $A \cap B$

- ١٣ - إذا كانت :-

$X = \{x | x^2 + x - 12 = 0\},$
$Y = \{y | y^2 = 4\}$

أوجد $X \cap Y$, $X \cup Y$

– ١٤ – إذا كانت :

$A = \{x | x^2 + x - 12 = 0\},$
$B = \{x | 2x - 6 = 0\}$

أوجد $A \cap B$

– ١٥ – إذا كانت :-

$x_1 = 2$, $x_2 = 5$, $x_3 = 6$, $x_4 = 3$, $x_5 = 4, x_6 = 9, x_7 = 11$ أوجد :-

(أ) $\sum_{i=1}^{7}(5x_i + 3)$ (ب) $\sum_{i=1}^{5}(x_i + c) + 2$ (جـ) $\sum_{i=1}^{4}(x_i + c)^2$

– ١٦ – إذا كانت :-

$x_1 = 8$, $x_2 = 9$, $x_3 = 6$, $x_4 = 3$, $x_5 = 4$, $x_6 = 7$ أوجد :-

(أ) $\sum_{i=1}^{6}(x_i^2 + x_i + 3)$ (ب) $\sum_{i=1}^{6}(x_i + x_i c^2)^3$ (جـ) $\sum_{i=1}^{6}(x_i^3 + 3i)$

– ١٧ – إذا كانت :- $x_1 = 7$, $x_2 = 5$, $x_3 = 2$, $x_4 = 10$, $x_5 = 8$ أوجد :-

(أ) $\sum_{i=1}^{5}(x_i^2 + 6x_i)$ (ب) $\sum_{i=1}^{5}(x_i + x_i^2 c)^2$ (جـ) $\sum_{i=1}^{5}(x_i + i)^2$

– ١٨ – إذا كانت :- $f(x) = 3x^2 + 4x + 2$ أوجد :-

(أ) $f(0)$ (ب) $f(-0.5)$ (جـ) $f(3)$

– ١٩ – يعطى حجم مستعمرة حشرية عند لحظة زمنية t (الزمن قيس بالأيام) بالعلاقة :-

$$f(t) = 5000 - \frac{1000}{1 + t^2}$$

أوجد :- (أ) $f(0)$ (ب) $f(2)$ (جـ) $f(5)$

– ٢٠ – يعطى حجم مزرعة بكتريا عند لحظة زمنية t (الزمن قيس بالساعات) بالعلاقة :-

$$f(t) = 2000 + 1000t - 1000t^2$$

أوجد :- (أ) $f(0)$ (ب) $f(2)$ (جـ) $f(5)$

الفصل الثالث

مقدمة في الاحتمال

Introduction to Probability

(٣-١) فراغ العينة والأحداث Sample Space and events

تُجرى الأبحاث في مجالات كثيرة، ففي مجال الطب قد يهتم باحث بدراسة تأثير دواء معين على الشفاء من مرضٍ ما، وفي مجال الاقتصاد قد يهتم باحث بدراسة أسعار ثلاث سلع مختلفة في فترات زمنية مختلفة، وفي مجال الزراعة قد يهتم باحث بدراسة تأثير سماد كيمائي على كمية المحصول. الطريق الوحيد للباحث للحصول على معلومات عن الظاهرة موضع الدراسة هو إجراء تجربة experiment وهى أى إجراء نحصل به على بيان (مشاهدة) سواء في الطبيعة أو في المعمل وهذا البيان قد يكون رقمي أو وصفي، كل مثال من الأمثلة التالية يمثل تجربة :

(ا) تسجيل درجة طالب.

(ب) قياس كمية المطر في يوم ما.

(جـ) فحص مصباح كهربائي وتسجيل عمره.

(د) ملاحظة وحدة منتجة وتدوين حالتها : سليمة أو تالفة.

(هـ) إلقاء عملة وملاحظة الوجه الذي يظهر.

نجد في معظم الحالات أن نتيجة التجربة تعتمد على عوامل الصدفة (عوامل خارجة عن إرادة الباحث أى في علم الله) ولايمكن التنبؤ بها بشيء من التأكيد، ولكن يمكن وصف فئة كل النتائج الممكنة لها قبل إجرائها.

تعريف : الفئة التي عناصرها تمثل جميع النواتج الممكنة لتجربة تسمى فراغ العينة.

مثال (٣-١) عند إلقاء زهرة نرد فإن فراغ العينة هو:-

$$S = \{1,2,3,4,5,6\}.$$

تعريف : يسمى أى عنصر في فراغ العينة نقطة عينة sample point .

مثال (٣-٢) عند إلقاء عملة ثلاث مرات فإن فراغ العينة هو :-

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

حيث H ترمز إلى ظهور صورة head و T ترمز إلى ظهور كتابة tail، فراغ العينة في هذه التجربة يتكون من 8 نقاط عينة.

تعريف : الحادثة event هى أى فئة جزئية من فراغ العينة.

في المثال (٣-٢) تمثل $A = \{HHH\}$ حادثة "ظهور ثلاث صور" ، أيضا $B = \{HHH, HHT, HTH, THH\}$ تمثل حادثة "ظهور صورتين على الأقل " وتعتبر B , A فئتان جزئيتان من فراغ العينة.

تعريف : إذا كانت الفئة الجزئية تحتوى على عنصر واحد فقط تسمى حادثة بسيطة simple event ، أما الحادثة المركبة compound event فهي التي تنتج من اتحاد أحداث بسيطة.

مثال (٣-٣) عند إلقاء عملتين مرة واحدة فإن فراغ العينة هو $S = \{HH, HT, TH, TT\}$. تمثل الحادثة $A = \{HH\}$ " ظهور صورتين " حادثة بسيطة بينما تمثل الحادثة $B = \{HH, HT, TH\}$ " ظهور صورة واحدة على الأقل" حادثة مركبة.

نقول أن الحادثة وقعت إذا ظهر أحد عناصرها عند إجراء التجربة. تمثل الفئة الخالية ϕ الحادثة المستحيلة الحدوث كما يمثل فراغ العينة S الحادثة المؤكدة الحدوث.

تعريف : يقال أن A , B حادثتان مانعتان (متنافيتان) exclusive events إذا كان وقوع إحداهما يمنع وقوع الآخر وفي هذه الحالة فإن $P(A \cap B) = \phi$.

فمثلا عند إلقاء عملة مرة واحدة فإن ظهور الصورة يمنع ظهور الكتابة وبالتالى فإن الحادثة التي تمثل ظهور الصورة والحادثة التي تمثل ظهور الكتابة حادثتان مانعتان.

مثال (٣-٤) عند إلقاء زهرة نرد مرة واحدة، بفرض أن الحادثة $A = \{2,4,6\}$ ظهور رقم زوجي والحادثة $B = \{1,3,5\}$ ظهور رقم فردى فإن $P(A \cap B) = \phi$. وعلى ذلك فإن A و B حادثتان مانعتان.

(٣-٢) طرق العد Counting Methods

يعتبر عنصر الصدفة المرتبط بظهور بعض الأحداث من المشاكل التي يقابلها الإحصائي ويحاول تقديرها عند إجراء التجربة ، تنتمي هذه المشاكل إلى فرع الاحتمال والذي سوف نتناوله في البند التالي. في كثير من الحالات نكون قادرين عل حل مشكلة الاحتمال عن طريق عد النواتج التي تنتمي إلى الحادثة محل السؤال وأيضا العدد الكلى لنواتج التجربة ، ولكن لبعض التجارب قد يكون عدد النواتج كبير جدا، وبالتالي قد يكون عمل قائمة تضمهم جميعا مهمة طويلة وصعبة. القاعدة الأساسية للعد في النظرية الآتية.

نظرية (٣-١) إذا أمكن إجراء عملية ما بطرق عددها n_1 وإذا أمكن إجراء عملية ثانية بطرق عددها n_2 و ... وإذا أمكن إجراء عملية k بطرق عددها n_k ، فإنه يمكن إجراء هذه العمليات معا بطرق عددها $n_1 \cdot n_2 \cdot \ldots \cdot n_k$.

مثال (٣-٥) توجد ثلاث طرق بين B , A وتوجد أربع طرق بين C , B بكم طريقة يمكن لسائق أن يصل من A إلى C .

الحل . عدد الطرق الممكنة للوصول من A إلى B هو $n_1 = 3$ وعدد الطرق الممكنة للوصول من B إلى C هو $n_2 = 4$ وعلى ذلك عدد الطرق الممكنة للوصول من A إلى C هو:-

$$n_1 \cdot n_2 = 3 \cdot 4 = 12.$$

مثال (٣-٦) يقدم مطعم 5 أصناف من اللحم و 3 أصناف من الحساء و 3 أصناف من السلطة و 4 أصناف من العصير. كم عدد الوجبات المختلفة التي يمكن تقديمها إذا علم أن الوجبة الواحدة تتكون من لحم وحساء وسلطة وعصير ؟

الحل . عدد الوجبات المختلفة التي يمكن تقديمها هي :-

$$n_1 \cdot n_2 \cdot n_3 \cdot n_4 = 5 \cdot 3 \cdot 3 \cdot 4 = 180.$$

عادة يكون الاهتمام بفراغ العينة الذي عناصره كل الترتيبات الممكنة لمجموعة من الأشياء ، على سبيل المثال، قد نهتم بمعرفة عدد الترتيبات الممكنة لجلوس ستة أشخاص على مائدة مستديرة ، الترتيبات المختلفة تسمى تباديل **Permutations** .

تعريف : التبديل هي ترتيب لكل أو جزء من فئة من الأشياء.

مثال (٣-٧) التباديل الممكنة لفئة الحروف a , b , c تكون :

$$abc \quad acb \quad bac \quad bca \quad cab \quad cba$$

أى يوجد ستة من التباديل الممكنة. باستخدام نظرية (٣-١) يمكن الوصول إلى نفس النتيجة بدون الحصول على قائمة بالترتيبات المختلفة. لدينا ثلاثة أماكن يمكن شغلها (ملئها) بالحروف a , b , c. سوف يكون لدينا ثلاثة اختيارات للمكان الأول، ولكن حيث أنه لا يمكن إحلال العنصر الأول فإنه يوجد اختيارين للمكان الثاني ويتبقى اختيار واحد للمكان الأخير حال إتمام اختيار الحرفين الأولين. وعليه فإن عدد التباديل هو $6 = 3 \cdot 2 \cdot 1 = 3!$. عموما فإن n من الأشياء المميزة يمكن ترتيبها بطرق عددها $n \cdot (n-1) \cdots 3 \cdot 2 \cdot 1$ وسوف نرمز لهذا الناتج بالرمز $n!$ الذي يسمى مضروب n ‏(**factorial n**) ، وعلى ذلك ثلاثة عناصر مختلفة يمكن ترتيبها بطرق عددها $3! = 3 \cdot 2 \cdot 1 = 6$. ويجب أن نعرف أن $1! = 1$ ، $0! = 1$.

نظرية (٣-٢) عدد تباديل n من الأشياء المميزة مأخوذة جميعا في نفس الوقت هو $n!$.

في بعض الأحيان قد يكون الاهتمام بعدد التباديل لأشياء مميزة عددها **n** مأخوذة **r** في كل مرة. فعلى سبيل المثال عدد تباديل الحروف a, b, c مأخوذة اثنين في كل مرة هو :

$$ab \quad ba \quad ac \quad ca \quad bc \quad cb.$$

باستخدام نظرية (٣-١) مرة أخرى يكون لدينا مكانين يمكن شغلهما بثلاثة اختيارات للمكان الأول واختيارين للمكان الثاني. وعلى ذلك يكون عدد الطرق لشغل المكانين

هو $3 \cdot 2 = 6$ ، سوف نرمز لعدد التباديل لأشياء عددها r من بين أشياء عددها n بالرمز $P(n, r)$.

نظرية (٣-٣) عدد تباديل n من الأشياء المميزة مأخوذة r في كل مرة هو :-

$$P(n,r) = n \cdot (n-1) \cdot \ldots \cdot (n-r+1) = \frac{n!}{(n-r)!}.$$

يراد أحيانا معرفة عدد تباديل مجموعة من الأشياء يكون بعضها متماثلا وتنتج الصيغة العامة لهذا العدد من النظرية التالية.

نظرية (٣-٤) عدد التباديل المختلفة لأشياء عددها n حيث n_1 من نوع و n_2 من نوع ثاني و...و n_k من النوع رقم k هو:-

$$\frac{n!}{n_1!\, n_2! \ldots n_k!}.$$

مثال (٣-٨) كم عدد التباديل المختلفة التي يمكن تكوينها من جميع حروف كل كلمة من الكلمات التالية على حدة ؟

(ا) منعم (ب) محمود

الحل . (ا) $\frac{4!}{2!1!1!} = \frac{4!}{2!} = 12$

لأنه يوجد 4 حروف أثنين منهم هو الحرف م .

(ب) $\frac{5!}{2!1!1!1!} = \frac{5!}{2!} = 60$

لأنه يوجد 5 حروف اثنين منها هو الحرف م.

مثال (٣-٩) كم عدد الطرق التي يمكن بها ترتيب 11 كتاباً على رف إذا كان 5 كتب منهم في التاريخ و 3 في الإحصاء و 3 في الرياضيات ؟

الحل . عدد الطرق هو :-

$$\frac{n!}{n_1!\, n_2!\, n_3!} = \frac{11!}{5!3!3!} = 9240.$$

عادةً يكون اهتمامنا بعدد الطرق لتجزئة فئة من الأشياء التي عددها n إلى فئات جزئية عددها r تسمى الخلايا cells . يتحقق التجزئة إذا كان التقاطع لأى زوج من الفئات الجزئية التي عددها r هو الفئة الخالية ϕ والاتحاد بين كل الفئات الجزئية يعطى الفئة الأصلية والترتيب للعناصر داخل الخلية ليس له أهمية. لتكن الفئة $\{a, b, c, d\}$ ، التجزينات الممكنة لهذه الفئة إلى خليتين بحيث تحتوى الخلية الأولى على ثلاثة عناصر والخلية الثانية تحتوى على عنصر واحد هي :

$$\{(a,b,c),d\},\{(a,b,d),c)\},\{(b,c,d),a\},\{a,c,d),b\}.$$

أى أن هناك 4 طرق لتجزئة الفئة المكونة من 4 عناصر إلى خليتين تحتوى على 3 عناصر في الخلية الأولى وعنصر واحد في الخلية الثانية. عدد التقسيمات للمثال السابق يمكن كتابتها على الشكل

$$\frac{4!}{3!1!}=4.$$

نظرية (٣-٥) عدد الطرق لتجزئة فئة من n من الأشياء إلى r من الخلايا بعناصر عددها n_1 في الخلية الأولى و n_2 من العناصر في الخلية الثانية و ... و n_r من العناصر في الخلية رقم r يكون :-

$$\frac{n!}{n_1!n_2!\ldots n_r!}.$$

حيث :-

$$n_1+n_2+\ldots+n_r=n$$

مثال (٣-١٠) بكم طريقة يمكن توزيع 10 كتب على 3 تلاميذ بحيث يتلقى التلميذ المتفوق 4 كتب وكل تلميذ آخر 3 كتب ؟

الحل . في هذا المثال يراد معرفة عدد التجزيئات لـ 10 كتب على ثلاث خلايا تحتوى على 4,3,3 من الكتب على التوالي . من النظرية السابقة عدد التجزيئات هو :-

$$\frac{10!}{4!3!3!}=4200.$$

مثال (٣-١١) بكم طريقة يمكن توزيع 9 أشخاص على 3 غرف في فندق حيث أن غرفتين من ذات سريرين وغرفة ذات 5 أسرة .

الحل . عدد الطرق هو :-

$$\frac{n!}{n_1!n_2!n_3!}=\frac{9!}{2!2!5!}=756.$$

في كثير من المشاكل يكون اهتمامنا بعدد الطرق لاختيار أشياء عددها r من بين أشياء مميزة عددها n ودون اعتبار لطريقة الترتيب. هذه الاختيارات تسمى التوافيق **combinations** . في الحقيقة التوفيقة **combination** هو تجزئة بخليتين، خلية تحتوى على r من الأشياء والخلية الأخرى تحتوى على $(n-r)$ من الأشياء الباقية وعدد هذه التوافيق يرمز له بالرمز $\binom{n}{r}$.

نظرية (٣-٦) عدد التوافيق لأشياء مميزة عددها n مأخوذة r كل مرة هو :

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

مثال (٣-١٢) كم عدد الطرق لاختيار 8 أشخاص لفريق كرة السلة من بين 11 ولدا ؟

الحل . عدد الطرق تكون :-

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{11!}{8!3!} = 165.$$

مثال (٣-١٣) كم عدد الطرق لاختيار عملتين من كيس يحتوى على دينار و ريال و درهم و ين و فرنك ؟

الحل . عدد الطرق هي :-

$$\binom{n}{r} = \binom{5}{2} = \frac{5!}{2!3!} = 10.$$

مثال (٣-١٤) بكم طريقة يمكن اختيار بعثة علمية تتكون من 3 رجال وسيدتين من بين 7 رجال و 5 سيدات .

الحل . يمكن اختيار 3 رجال من بين 7 رجال بطرق عددها $\binom{7}{3} = 35$ ويمكن اختيار السيدتين من بين 5 سيدات بطرق عددها $\binom{5}{2} = 10$ وبذلك يكون اختيار البعثة بطرق عددها:-

$$\binom{7}{3}\binom{5}{2} = 350.$$

(٣-٣) الاحتمال **Probability**

عادة، يكون هدف الإحصائي الوصول إلى استنتاجات أو استدلالات من التجارب التي تشتمل على أحداث غير مؤكدة. ولكي تكون استنتاجاته واستدلالاته دقيقه فإن فهمه لنظرية الاحتمالات يكون ضروري. ماذا نعني عندما نقول مثلا أنه من المحتمل أن تمطر السماء بعد ظهر اليوم أو أن الاحتمال ضئيل في أن ينجح طالب معين في الامتحان ...الخ. في كل حالة نعبر عن نتيجة غير متأكدين منها ولكن من معلوماتنا السابقة أو فهمنا لطبيعة التجربة يكون لدينا درجة من الثقة في صحة هذه الجملة .

تمدنا نظرية الاحتمالات بفئة من الأرقام تسمى الأوزان **weights** تتراوح من الصفر إلى الواحد الصحيح والتي تمكننا من تقدير لإمكانية (فرصة) وقوع الأحداث التي تنتج من تجارب إحصائية. لكل نقطة في فضاء العينة نعين وزن بحيث يكون مجموع الأوزان يساوى الواحد الصحيح. إذا كان لدينا سبب لكي نعتقد أن هناك إمكانية كبيرة لوقوع نقطة في فضاء العينة

فإننا نعين لها رقماً قريباً من الواحد الصحيح. ومن ناحية أخرى يعين وزن قريب من الصفر لنقاط العينة التي إمكانية وقوعها ضئيل • للنقاط خارج نطاق العينة، أى النقاط التي يستحيل حدوثها فإننا نعين لها الرقم صفر وتسمى الأحداث المستحيلة الحدوث • على سبيل المثال احتمال ظهور الرقم 7 عند إلقاء زهرة نرد مرة واحدة يساوى صفر. لإيجاد احتمال أى حدث A فإننا نجمع كل الأوزان المعينة لنقاط العينة في A. هذا المجموع يسمى المقياس **measure** للحادثة A أو احتمال A ويرمز له بالرمز $P(A)$ • الطرق المختلفة لقياس الاحتمالات تمثل مفاهيم مختلفة. في هذا البند سوف نناقش ثلاثة مفاهيم مختلفة لقياس الاحتمالات وهي : المفهوم القديم (المفهوم الكلاسيكي **classical concept**) و مفهوم التكرار النسبي **relative frequency concept** والمفهوم الشخصي **subject concept** .

(٣-٣-١) المفهوم القديم (المفهوم الكلاسيكي)**Classical Concept**

تبعا لهذا المفهوم تحدد أرقام الاحتمالات أو يمكن تقديرها قَبْلِي **a priori** (قبل الحقيقة **before fact**) • وعلى ذلك، الاحتمال بالضبط **exact probability** أن حادثة ما تقع تحدد قبل وقوع الحادثة. لذلك يسمى الاحتمال المقدر بناء على هذا المفهوم بالاحتمال القبلي **a priori probability** . المفهوم القديم للاحتمال مبنى على أساس أنه إذا كانت تجربة تحتوى على M من النقاط، أى أن عدد النواتج الممكنة لتجربة ما هو M وكانت هذه النواتج متساوية في إمكانية الحدوث وإذا احتوت الحادثة A على عدد m من النقاط فإن احتمال الحادثة هو :

$$P(A) = \frac{m}{M}$$

مثال (٣-١٥) أوجد احتمال ظهور عددين مجموعهما ثمانية عند إلقاء زهرتي نرد مرة واحدة.

الحل . فراغ العينة هو فئة الأزواج المرتبة الآتية :-

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

وكل زوج مرتب من هذه الفئة يمثل أحد نتائج التجربة. فمثلا العنصر (1,6) يمثل ظهور العدد 1 على الزهرة الأولى والعدد 6 على الزهرة الثانية. فراغ العينة يحتوى على ستة وثلاثين زوجاً مرتبا. المجموع ثمانية على النردين سينتج إذا ظهر أى زوج من الأزواج الخمسة التالية :-

(3,5) , (5,3) , (4,4) , (6,2) , (2,6)

أى أن "ظهور عددين مجموعهم يساوى ثمانية " ينتج من 5 نقاط عينة ، وعلى ذلك :-

$$P(A) = \frac{m}{M} = \frac{5}{36}.$$

مثال (٣-١٦) أسرة لديها 3 أطفال. ما احتمال أن يكون جميعهم أولاد (بفرض أن كل طفل له نفس الاحتمال أن يكون ولدا أو بنتا).

الحل . إذا رمزنا للولد بالرمز B وللبنت بالرمز G فإن فراغ العينة هو :-

$$S = \{BBB, BBG, BGB, GBB, GBG, BGG, GGB, GGG\}$$

الحادثة " جميع الأطفال أولاد " هو $A = \{BBB\}$ ، أي أن $M = 8, m = 1$ وبالتالى فإن:-

$$P(A) = \frac{m}{M} = \frac{1}{8}.$$

(٣-٣-٢) مفهوم التكرار النسبي **Relative Frequency Concept**

يشترط هذا المفهوم إجراء التجربة عدد كبير من المرات ومعرفة نتائجها وبعد ذلك قياس الاحتمال. فإذا كانت N تمثل عدد المرات (المحاولات **trails**) التي أجريت بها تجربة ما تحت نفس الظروف و n تمثل عدد مرات (التكرار) ظهور الحادثة A خلال N من المرات التي كررت فيها التجربة فإن احتمال وقوع الحادثة A هو :-

$$P(A) = \lim_{N \to \infty} \frac{n}{N}$$

حيث $\frac{n}{N}$ هو التكرار النسبي للحادثة A في هذه التجارب التي عددها N. عادة تكون قيم التكرار النسبي غريبة الأطوار للقيم الصغيرة من N ولكن عندما تزيد قيمة N، فقد أوضحت الخبرة أن النسبة $\frac{n}{N}$ تكتسب بعض الانتظام الإحصائي وتستقر حول قيمة ثابتة هي $P(A)$ ولذلك عرف الاحتمال بأنه نهاية النسبة $\frac{n}{N}$ عندما يزداد عدد المحاولات أو التجارب ويؤول إلى ما لانهاية. وحيث أن عدد المحاولات نادرا ما يؤول إلى ما لانهاية فإننا نستخدم التكرار النسبي كتقدير لقيمة الاحتمال. الاحتمال المبنى على هذا المفهوم يقدر بعدى **a posteriori** (بعد الحقيقة **after fact**) ويعتمد على البيانات الملاحظة، لذلك يسمى في بعض الأحيان الاحتمال التجريبي **experiment probability** .

مثال (٣-١٧) في مصنع لإطارات السيارات تبين أن كل 100000إطار منتج يكون من بينها 300 إطار تالف. فما احتمال اختيار إطار تالف؟

الحل . عدد الإطارات N=100000 عدد الإطارات التالفة 300 = n ، وعلى ذلك احتمال اختيار إطار تالف هو :-

$$P(A) \approx \frac{300}{100000} = 0.003.$$

(٣-٣-٣) المفهوم الشخصي **Subject Probability**

تبعا لهذا المفهوم، الاحتمال هو درجة الثقة **degree of confidence** في وقوع حادثة ما والمقررة من شخص ما بناء على دليل متوفر لديه **evidence available** ، هذا الدليل قد يكون أي معلومات كمية أو غير كمية. على سبيل المثال قد يحدد الشخص القائم على المشتريات في شركة ما الاحتمال 0.25 للحادثة أن شحنة ما تحتوى على الأكثر 2% وحدات تالفة ، أيضا قد يصرح المدير الأول في شركة بأن احتمالات زيادة ميزانية الشركة أو انخفاضها أو بقائها على حالها هو 0.04 و 0.13 و 0.83 على التوالي. ويجب ملاحظة أن الاحتمال المقدر لحادثة ما بناء على هذا المفهوم يختلف من شخص إلى آخر وذلك لعوامل كثيرة منها الخبرة ،

(٣-٣-٤) الخواص المميزة لقيم الاحتمال

Characteristics of Probability Numbers

إذا كان S فراغ العينة المرافق لتجربة وإذا كانت $A_1, A_2, \ldots$ تمثل كل الأحداث الممكنة فإن قيم الاحتمال المقدرة للأحداث السابقة لابد أن تتوافر فيها الشروط الآتية :-

(ا) يرافق كل حادثة A عدد معين $P(A)$ يسمى احتمال A ويحقق $P(A) \geq 0$.

(ب) احتمال وقوع حادثة مؤكدة يساوى الواحد الصحيح، أي أن $P(S) = 1$

(جـ) إذا كانت $A_1, A_2, A_3 \ldots$ عدد لانهائي من الحوادث المانعة بالتبادل أي $A_i \cap A_j = \phi \; , \; i \neq j$ فإن :-

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots.$$

ويمكن إثبات أنه إذا كانت $A_1, A_2, \ldots, A_n$ تمثل n حادثة مانعة بالتبادل فإن :-

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

أيضا إذا كانت $A_1, A_2, \ldots A_n$ تمثل تجزئة لفراغ العينة S فإن :

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) = 1.$$

مثال (٣-١٨) ما هو احتمال الحصول على مجموع 5 أو 12 عند إلقاء زهرتا نرد مرة واحدة.

الحل . بفرض أن A حادثة "ظهور 5 " و B حادثة "ظهور 12 " . المجموع 5 يحدث من 4 نقط عينة والمجموع 12 يحدث من نقطة واحدة. وحيث أن كل النتائج في فضاء العينة متساوية في إمكانية الحدوث فإن $P(A) = \frac{4}{36}$ ، $P(B) = \frac{1}{36}$. الحادثتان A ، B مانعتان وعلى ذلك :-

$$P(A \cup B) = P(A) + P(B)$$
$$= \frac{4}{36} + \frac{1}{36} = \frac{5}{36}.$$

(٣-٤) بعض قوانين الاحتمال Some Probability Laws

عادة يكون من السهل حساب احتمال حادثة ما من الاحتمالات المعروفة للأحداث الأخرى وهذا يكون صحيح إذا أمكن تمثيل الحادثة كاتحاد لحادثتين أخرتين أو مكملة لحادثة. في هذا البند سوف نعرض بعض القوانين التي تسهل عملية حساب الاحتمالات.

نظرية (٣-٧) لأى حادثتين A , B فإن :-

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

البرهان :-

كما يتضح من شكل (٣-١) أن $I = A^c \cap B$ ، $II = A \cap B^c$ ، $III = A \cap B$

يمكن تمثيل كل من الحادثتين $A \cup B$ و A كاتحاد لحادثتين مانعتان كما يلي :-

$$A \cup B = (A \cap B^c) \cup B,$$
$$A = (A \cap B) \cup (A \cap B^c).$$

حيث :

$$(A \cap B^c) \cap B = \phi,$$
$$(A \cap B) \cap (A \cap B^c) = \phi.$$

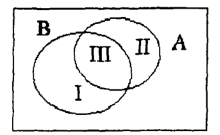
وعلى ذلك :-

$$P(A \cup B) = P(A \cap B^c) + P(B),$$
$$P(A) = P(A \cap B) + P(A \cap B^c).$$

وبذلك نصل إلى :-

$$P(A \cup B) = P(A \cap B^c) + P(B)$$
$$= [P(A) - P(A \cap B)] + P(B)$$

$$= P(A) + P(B) - P(A \cap B).$$



شكل (٣-١)

مثال (٣-١٩) وجد أن 40% من المرضى الذين تم فحصهم في العيادة المحلية يعـــانون مــن ارتفاع في ضغط الدم ، وأن 30% يعانون من زيادة الوزن وأن 10% يعانون من كليـــهما. إذا اختير مريضا عشوائيا ، فما هو احتمال أن يعاني من إحدى هاتين الحالتين على الأقل ؟

الحل، بفرض A حادثة " ارتفاع ضغط الدم " و B حادثة " الوزن زائد " و $A \cup B$ حادثة "المعاناة من إحدى هاتين الحالتين " . من نظرية (٣-٧) :-

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.4 + 0.3 - 0.1 = 0.6.$$

نظريه (٣-٨) إذا كانت A حادثة وكانت A^c الحادثة المكملة لها فإن :-

$$P(A) = 1 - P(A^c).$$

البرهان :-

نعلـــم أن A^c هـــي المكملـــة للفئـــة A وعلـــى ذلـــــك $S = A \cup A^c$. وحيـــث أن : $A \cap A^c = \phi$ لأن A , A^c حادثتين مانعتان ، فإن :-

$$P(S) = P(A \cup A^c)$$

إذن :

$$P(S) = P(A) + P(A^c)$$
$$1 = P(A) + P(A^c).$$

وعلى ذلك :-

$$P(A) = 1 - P(A^c).$$

مثال (٣-٢٠) إذا ألقيت عملة 7 مرات أوجد احتمال ظهور صورة مرة على الأقل.

الحل• بفرض A الحادثة " ظهور صورة مرة على الأقل " • عدد نقط العينة في S هو 2^7 (من نظرية (٣-١))• بفرض أن الحادثة A^c تمثل عـــدم ظـــهور أي صـــورة وحيـــث أن الأحداث متساوية في إمكانية الحدوث فإن :

$$P(A^c) = \frac{1}{2^7} = 0.0078125.$$

أي أن :-

$$P(A) = 1 - P(A^c)$$
$$= 1 - 0.0078125 = 0.9921875.$$

(٣-٥) الاحتمال الشرطي **Conditional Probability**

في بعض التجارب يتأثر الاحتمال الذي يخصص لحادثة ما (لتكن A) بالمعلومات عـــن حدوث أو عدم حدوث حادثة أخرى ولتكن B • في هذه الحالة سوف نســـتخدم العبـــارة : احتمال وقوع حادثة A بشرط وقوع حادثة B والذي يسمى الاحتمال الشرطي ويرمـــز لـــه بالرمز $P(A|B)$ ويقرأ " احتمال وقوع A الشرط وقوع B "• للتســـهيل بفـــرض أن A حادثة الحصول على رقم 3 عند إلقاء زهرة نرد متزنة مرة واحـــدة. بالاعتمـــاد علـــى فضـــاء العينة $S = \{1,2,3,4,5,6\}$ فإن احتمال الحصول على الرقم 3 عند إلقاء زهـــرة نـــرد مـــرة واحدة هو $\frac{1}{6}$ • الآن بفرض أن الشخص الذي قام بإلقاء زهرة النرد أفادنا بأن النتيجة الـــتي حصل عليها كانت رقم فردى • لنستعرض آثار هذه المعلومات التي توفرت مسبقاً على احتمـــال الحادثة A والتي حسبناها مسبقا. الآن في ظل هذه الحقيقة لم تعد النتائج الممكنـــة ســـتة وإنمـــا أصبحت ثلاثة فقط فهي إما 1 أو 3 أو5. أما النتائج2 أو4 أو 6 فأصبحت مستحيلة. وعلى ذلك فإن احتمال الحادثة A منسوبا إلى الفراغ الجديد المختزل $B = \{1,3,5\}$ هو :

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{1}{3}.$$

حيث $n(B), n(A \cap B)$ تمثل عدد العناصر التابعة للحادثتين $B, A \cap B$ على التوالى•

مثال (٣-٢١) ألقيت عملة ثلاث مرات فإذا علم أن الوجه الظاهر في الرمية الأولى والثانيـــة كتابة ما هو احتمال ظهور كتابة في الرمية الثالثة ؟

الحل • فراغ العينة هو :-

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

بفرض أن الحادثة $A = \{HHT, THT, HTT, TTT\}$ "ظهور كتابة في الرمية الثالثـــة" و الحادثـــة $B = \{TTH, TTT\}$ " ظـــــهور كتابـــة في الرميـــة الأولى والثانيـــة "• الحادثة $A \cap B = \{TTT\}$ تحتوى على نقطة واحدة والحادثة B تحتـــوى علـــى نقطتـــين. باستخدام فراغ العينة المختزل B ، وإذا كانت $n(B), n(A \cap B)$ تمثل عدد العناصر التابعة للحادثتين $A \cap B$, B على التوالى• وعلى ذلك يكون الاحتمال المطلوب هو :

$$P(A \mid B) = \frac{n(A \cap B)}{n(B)} = \frac{1}{2}.$$

مثال (٣-٢٢) وعاء يحتوى على 150 وحدة بعضها منتج من المصنع 1 والباقي من المصنع 2، بعض الوحدات سليمة وبعضها تالفة• • فإذا اختيرت وحـــدة عشـــوائية مـــن الوعـــاء، ليكن A الحادثة "الوحدة تالفة" وعلى ذلك، A^c الحادثة "الوحدة سليمة" • أيضا ليكـــن B الحادثة "الوحدة من إنتاج المصنع 1" ، B^c الحادثة "الوحدة من إنتاج المصنع 2 " . الجـــدول التالي يمثل عدد الوحدات السليمة والتالفة المنتجة من المصنعين . الآن بفرض أن كـــل وحـــدة مرقمة بعلامة توضح المصنع المنتج أوجد $P(A|B)$ •

	B	B^c	المجموع
A	50	5	55
A^c	75	20	95
المجموع	125	25	150

الحل • يمكن حساب احتمال أن الوحدة تالفة بشرط وقوع الحادثة B وباستخدام الفـــراغ المختزل B كالتالي :-

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{50}{125} = \frac{2}{5}.$$

أيضا يمكن كتابة $P(A \mid B)$ على الشكل :

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{n(A \cap B)/n(S)}{n(B)/n(S)} = \frac{P(A \cap B)}{P(B)},$$

حيث $P(B)$, $P(A \cap B)$ يتم الحصول عليهما من فراغ العينة الأصلي S • للتحقق من النتيجة فإن :-

$$P(B) = \frac{125}{150} \quad ,$$
$$P(A \cap B) = \frac{50}{150}.$$

وعلى ذلك :-

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{50}{125} = \frac{2}{5}.$$

تعريف : الاحتمال الشرطي للحادثة A شرط B يمثل بالصيغة $P(A|B)$ و يعرف بالمعادلة :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \; , \; P(B) \neq 0.$$

يعتبر هذا التعريف عام ولا يعتمد على فراغ عينة يحتوى على أحداث متساوية في إمكانية الحدوث كما في الأمثلة السابقة وبنفس الشكل يمكن القول أن :

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \; , \; P(A) \neq 0.$$

مثال (٣- ٢٣) في إحدى الكليات نجح 75% من الطلبة في امتحان الرياضيات ونجح 85% من الطلبة في امتحان الإحصاء ونجح 10% في الرياضيات والإحصاء • اختير أحد الطلبة بطريقة عشوائية • المطلوب :-

(ا) إذا كان ناجحا في الإحصاء ما هو احتمال أن يكون ناجحا في الرياضيات •

(ب) إذا كان ناجحا في الرياضيات فما هو احتمال أن يكون ناجحا في الإحصاء •

الحل• بفرض أن A حادثة " النجاح في الرياضيات" و B حادثة "النجاح في الإحصاء"

وعلى ذلك :-

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.1}{0.85} = \frac{10}{85} = \frac{2}{17} \qquad (ا)$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.1}{0.75} = \frac{10}{75} = \frac{2}{15} \qquad (ب)$$

نظرية (٣-٩) إذا وقعت حادثة ما A في تجربة ما، يتبعها حادثة B فإن :-

$$P(A \cap B) = P(A)P(B|A).$$

وعلى ذلك احتمال وقوع A , B في ترتيب هو احتمال أن تقع A أولا مضروبا في احتمال وقوع B، شرط أن A وقعت • كما يمكن أن يكون:-

$$P(A \cap B) = P(B)P(A|B).$$

وهذا يتوقف على أى الحادثتين قد تقع أولا ٠

مثال (٣-٢٤) كيس يحتوى على 4 كرات بيضاء و7 حمراء فإذا أختار شخص كرتين من الكيس اختيارا عشوائيا فما احتمال أن تكون الأولى بيضاء والثانية حمراء ؟ (إذا كان السحب بدون إرجاع) ٠

الحل ٠ بفرض أن A الحادثة "الكرة الأولى بيضاء" و B الحادثة "الكرة الثانية حمراء" وعلى ذلك احتمال أن تكون الكرة الأولى بيضاء والثانية حمراء هو :-

$$P(A \cap B) = P(A)P(B|A) = (\frac{4}{11})\ (\frac{7}{10}) = \frac{14}{55}.$$

نظرية (٣-١٠) في أي تجربة إذا وقعت الحادثة A_1 ، يتبعها الحادثة A_2 ، يتبعها الحادثة A_3 ، وهكذا ، فإن :-

$$P(A_1 \cap A_2 \cap A_3 \cap ...) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2)...$$

في بعض الأحيان ، احتمال وقوع حادثة ما A لا يتأثر ولا يعتمد على وقوع أو عدم وقوع حادثة أخرى B ٠ بعبارة أخرى وفي هذه الحالة يقال أن A مستقلة عن B وعلى ذلك $P(A|B) = P(A)$ وفى هذه الحالة فإن :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A).$$

ومنها :

$$P(A \cap B) = P(A)P(B).$$

إذا كانت A مستقلة عن B فإن B تكون مستقلة عن A لأن :-

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B).$$

ومنها :

$$P(A \cap B) = P(A)P(B).$$

تعريف : يقال أن الحادثتين A ، B مستقلتين **independent** ، إذا وفقط إذا :-

$$P(A \cap B) = P(A)P(B).$$

مثال (٣-٢٥) في المثال السابق إذا كان السحب بإرجاع المطلوب إيجاد احتمال أن تكون الكرة الأولى بيضاء والثانية حمراء؟

الحل، بفرض أن A الحادثة " الكرة الأولى بيضاء " ، B الحادثة " الكـــرة الثانيــة حمـــراء " الحادثتين A و B مستقلتين وعلى ذلك احتمال أن تكون الكرة الأولي بيضاء والثانية حمــــراء هو:-

$$P(A \cap B) = P(A)P(B) = (\frac{4}{11})\ (\frac{7}{11}) = \frac{28}{121}.$$

مثال (٣-٢٦) ألقيت عملة وزهرة نرد معا، ما هو احتمال ظهور الصورة على العملة والرقـــم 6 على زهرة النرد ؟

الحل، بفرض أن الحادثة A "ظهور الصورة على العملة " والحادثة B ظـــهور رقـــم 6 علـــى النرد، ولأن الحادثتين مستقلتين فإن :-

$$P(A \cap B) = P(A)P(B)$$
$$= (\frac{1}{2})(\frac{1}{6}) = \frac{1}{12}.$$

مثال (٣-٢٧) أُلقيت زهرتيّ نرد مرة واحدة ما هو احتمال ظهور الرقم 5 علـــى الزهـــرة الأولى والرقم 2 على الزهرة الثانية ؟

الحل، بفرض أن A الحادثة ظهور الرقم 5 على الزهرة الأولى، B الحادثة ظهور الرقم 2 علـــى الزهرة الثانية ،بما أن الحادثتين مستقلتين فإن:-

$$P(A \cap B) = P(A)P(B)$$
$$= (\frac{1}{6})(\frac{1}{6}) = \frac{1}{36}.$$

مثال (٣-٢٨) أُلقيت زهرتيّ نرد مرتين، ما هو احتمال أن مجموع الوجهين 5 في رمية و7 في الرمية الأخرى ؟

الحل، بفرض أن A_1 ,A_2 ,B_1 , B_2 على التوالى ظهور مجموع 5 في الرمية الأولى و مجموع 5 في الرمية الثانية و ظهور مجموع 7 في الرمية الأولى و مجموع 7 في الرمية الثانية (تمثل أحداث مستقلة) ، اهتمامنـــا ســـوف يكـــون في حســـاب احتمـــال الاتحـــاد لحـــادثتين مـــانعتين $A_1 \cap B_2$ ، $B_1 \cap A_2$ ، وعلى ذلك :-

$$P[(A_1 \cap B_2) \cup (B_1 \cap A_2)] = P(A_1 \cap B_2) + P(B_1 \cap A_2)$$
$$= P(A_1)P(B_2) + P(B_1)P(A_2)$$
$$= (\frac{4}{36})(\frac{6}{36}) + (\frac{6}{36})(\frac{4}{36}) = 0.037037.$$

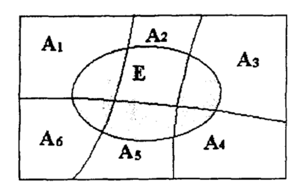
(٣-٦) الاحتمال الكلي وقاعدة بييز

Total Probability and Bayes` Rule

بفرض أن الأحداث $A_1, A_2, ..., A_n$ تمثل تجزيئا لفراغ العينة ومانعة لبعضها البعـــض واتحادهم هو S (أحداث مانعة وشاملة **mutually exclusive and exhaustive**)

كما في شكل (٣-٢) حيث **n=6** ، بفرض أن E أى حادثة أخرى فإن :-

$$E = S \cap E = (A_1 \cup A_2 \cup ... \cup A_n) \cap E$$
$$= (A_1 \cap E) \cup (A_2 \cap E) \cup ... \cup (A_n \cap E)$$



شكل (٣-٢)

الصيغة السابقة مفيدة في النظرية الآتية :-

نظرية (٣-١١) (نظرية الاحتمال الكلى **total probability**)

بفرض أن $A_1, A_2, \ldots, A_n$ تمثل n حادثة مانعة وشاملة، وعلى ذلك لأى حادثة E فإن :-

$$P(E) = \sum_{i=1}^{n} P(A_i) P(E|A_i).$$

البرهان :-

الأحداث $A_1 \cap E, A_2 \cap E, ..., A_n \cap E$ مانعة بالتبادل، وعلى ذلك :-

$$P(E) = \sum_{i=1}^{n} P(A_i \cap E),$$

وبتطبيق نظرية (٣-٩) على كل حد نحصل على برهان النظرية .

مثال (٣-٢٩) تنتج ثلاث ماكينات A , B , C 40% , 35% , 25% على التـــوالي من الإنتاج الكلى لمصنع، ونسبة الإنتاج السليم لهذه الماكينات هـــى 98% , 96% , 94% فإذا اختيرت وحدة بطريقة عشوائية، ما هو الاحتمال أن تكون سليمة ؟

الحل • بفرض أن الحادثة E "الوحدة المختارة سليمة" ، والحادثة A_1 "الوحدة من إنتاج الماكينة **A**" والحادثة A_2 "الوحدة من إنتاج الماكينة **B**" والحادثة A_3 "الوحدة من إنتاج الماكينة **C**" وعلى ذلك :-

$$P(E) = \sum_{i=1}^{3} P(A_i)P(E \mid A_i)$$

$$= (0.4)(0.98) + (0.35)(0.96) + (0.25)(0.94)) = 0.9630 .$$

نظرية (٣-١٢) نظرية بيز **Bayes' Theorem**

إذا كانت $A_1, A_2, \ldots, A_n$ تمثل n حادثة مانعة وشاملة وكان ظهور إحداهما ينتج عنه ظهور حادثة أخرى E (أى أن E تقع إذا وقعت واحدة من الحوادث المانعة) فإن :-

$$P(A_k \mid E) = \frac{P(A_k)P(E \mid A_k)}{\sum_{i=1}^{n} P(A_i)P(E \mid A_i)}, k = 1,2,\ldots,n.$$

البرهان :-

نعلم من النظرية السابقة أن :-

$$P(E) = \sum_{i=1}^{n} P(A_i)P(E|A_i).$$

وحيث أن :-

$$P(A_k|E) = \frac{P(A_k \cap E)}{P(E)} = \frac{P(A_k)P(E|A_k)}{P(E)}.$$

وبالتعويض عن $P(E)$ بقيمتها من نظرية (٣-١١) يتم البرهان •

مثال (٣-٣٠) تُنتج إحدى شركات المشروبات نوع معين من العصائر • يستمر الإنتاج خلال ورديتين بحيث أن 70% من الإنتاج اليومي من الوردية الأولى • من دراسة المنتج وجد أن نسبة العبوات السليمة من إنتاج الوردية الأولى 95% ونسبة العبوات السليمة من إنتاج الوردية الثانية 97% • فإذا سُحبت إحدى العبوات عشوائيا وكانت سليمة ما هو احتمال أن تكون من إنتاج الوردية الثانية ؟

الحل • بفرض أن E الحادثة "العبوة سليمة" و A_1 الحادثة "العبوة من إنتاج الوردية الأولى" و A_2 الحادثة "العبوة من إنتاج الوردية الثانية " وعلى ذلك الاحتمال المطلوب هو :-

$$P(A_2 \mid E) = \frac{P(A_2)P(E \mid A_2)}{P(A_1)P(E \mid A_1) + P(A_2)P(E \mid A_2)}$$

$$= \frac{(0.3)(0.97)}{(0.7)(0.95) + (0.3)(0.97)} = 0.3043933.$$

تماريـــــــن

- ١ - أُلقى زوج من زهرتي النرد مرة واحدة أذكر الحادثة :-

(ا) مجموع الوجهين الظاهرين يساوى 9 .

(ب) مجموع الوجهين الظاهرين إما 4 أو 5 .

- ٢ - أُلقيت عملتين مرة واحدة أذكر الحادثة :-

(ا) ظهور كتابة واحدة.

(ب) ظهور كتابة واحدة على الأقل.

- ٣ - في تجربة لاختيار ثلاث وحدات من مصنع وملاحظة ما إذا كانت الوحدة سليمة أو تالفة (يرمز للتالفة D والسليمة D') أذكر :-

(ا) فضاء العينة.

(ب) الحادثة عدم ظهور وحدات تالفة.

(جـ) كيف يمكن تعريف الحادثة ؟

$$A = \{(DDD'), (DD'D), (D'DD)\}$$

- ٤ - اختير أربعة أشخاص عشوائيا لاختبار تفضيل أو عدم تفضيل لنوع معين من القهوة حيــث يعطى 1 للتفضيل و 0 لعدم التفضيل أذكر :-

(ا) فضاء العينة.

(ب) الحادثة ثلاث أشخاص على الأقل يفضلون.

- ٥ - قام مسئول بمراقبة الجودة في مصنع لإنتاج أسماك السلامون باختبار كل صندوق منتج وأخذ عينة والاستمرار في الاختبار حتى ظهور صندوق تالف. أذكر فضاء العينة لعملية الاختبار مع العلم أن Y ترمز للصندوق السليم و N ترمز للصندوق التالف.

- ٦ - قام باحث متخصص في التسويق بتصنيف العملاء إلى ثلاث مجموعات حســب الدخــل: منخفض 0 ومتوسط 1 وعالي 2 . كما قام بتصنيفهم تبعاً لخاصية أخرى وهى القوة الشرائية إلى (لا يشترى 0) و (يشترى ولو مرة واحدة في الشهر 1). عرف فضاء العينة.

- ٧ - بفرض عدم السماح بالتكرار (ا) كم عددا مكون من ثلاث أرقام يمكن تركيبة من الأرقام التالية 8,7,3,2,1 ؟ (ب) كم عدداً منهم زوجياً ؟ (جـ) كم عدداً منهم فردياً ؟

- ٨ - بكم طريقة يمكن أن يجلس أربعة طلاب من قسم الكيمياء وأربعة طلاب من قسم النبات وثلاثة طلاب من قسم الرياضيات وطالبان من قسم الحيوان في صف بحيث يجلس الأشخاص ذو التخصصات الواحدة معا ؟

- ٩ - حل نفس التمرين السابق إذا جلس الجميع حول مائدة مستديرة .

- ١٠ - كم عدد الطرق لاختيار ثلاثة عملات من صندوق يحتوى على جنيه و ريال و دينار و ين و فرنك ؟

- ١١ - كم عدد الطرق لاختيار ثمانية أشخاص لفريق كرة القدم من 14 شخصا ؟

- ١٢ - صنعت زهرة نرد بحيث يكون احتمال ظهور الرقم واحد ثلاثة أضعاف أي رقم آخر، بينما كل الوجوه الأخرى لها نفس الفرصة في الظهور . ما هو احتمال ظهور الرقم اثنين عند إلقاء النرد مرة واحدة ؟ وما هو احتمال ظهور الرقم واحد ؟

- ١٣ - مطلوب من طالب دراسة مادة في العلوم ومادة في الرياضيات ومادة في الاجتماع . ما هو عدد الطرق لاختيار هذه المواد من بين 3 مواد في العلوم و 4 في الاجتماع و مادتين في الرياضيات ؟

- ١٤ - ما عدد الطرق الممكنة لشخص داخل محل ملابس لاختيار رابطة عنق و قميص وذلك إذا توافر له 4 أربطة عنق و 5 قمصان في المحل ؟

- ١٥ - بكم طريقة يمكن زراعة 8 شجرات على شكل دائرة ؟

- ١٦ - كم عددا مكون من ثلاثة أرقام التي يمكن تكوينها من الأعداد 0,1,2,3,4,5 ؟ وإذا كان كل رقم يظهر مرة واحدة (ا) كم عدد الأرقام الفردية ؟ (ب) كم عدد الأرقام الزوجية ؟

- ١٧ - إذا لعب فريق كرة القدم ثمانية مباريات خلال الموسم . بكم طريقة يستطيع الفريق في نهاية الموسم أن يكسب 4 ويفقد 3 ويتعادل 1 ؟

- ١٨ - بكم طريقة يمكن الإجابة 10 على أسئلة من نوع صح وخطأ ؟

- ١٩ - أعطى امتحان في مادة الإحصاء لطالب . يتكون الامتحان من 9 أسئلة منهم 6 أسئلة اختياري وثلاثة إجباري . فإذا كان المطلوب منه الإجابة على ستة أسئلة . بكم طريقة يمكن للطالب اختيار الأسئلة التي يرغب الإجابة عليها ؟

- ٢٠ - بكم طريقة يمكن لمدرس أن يختار طالبا من بين سبعة طلاب ؟

- ٢١ - أوجد عدد الطرق التي يمكن بها تصنيف 5 كتب من الحجم الكبير و 4 من الحجم المتوسط و 3 من الحجم الصغير على إحدى الرفوف بشرط أن تكون جميع الكتب ذات الحجم الواحد مصفوفة معا .

– ٢٢ – ثلاث أجزاء من كتاب موضوعة على رف ما احتمال (ا) الأجزاء في وضعها الصحيح ؟ (ب) الجزء الثاني في المكان الأول ؟

– ٢٣ – اختيرت ثلاثة كتب عشوائيا من رف يحتوى على 5 كتب في التاريخ و3 كتب في العلوم وقاموس ما هو احتمال (ا) القاموس هو المختار ؟ (ب) كتابين في العلوم و واحد في التاريخ هما المختارتان ؟

– ٢٤ – في مدينة ما احتمال أن أسرة تشترى تلفزيون هو 0.8 واحتمال أن تشترى غسالة ملابس هو 0.5 واحتمال أن تشترى الاثنين معا هو 0.45 ، ما هو احتمال أن تشترى الأسرة واحد من الاثنين على الأقل ؟

– ٢٥ – احتمال أن تمطر السماء في بلد ما في 4 يوليو هو 0.1 واحتمال حدوث رعد هو 0.5 واحتمال حدوث مطر أو رعد هو 0.3 ما هو احتمال حدوث مطر ورعد في نفس اليوم ؟

– ٢٦ – لاعب كرة يكسب 50% من مبارياته . ما هو احتمال أن يكسب بالضبط 3 من الأربع مباريات القادمة ؟

– ٢٧ – وعاء يحتوى على 10 وحدات منهم 3 تالفين سحبت وحدتين من الوعاء الواحدة تلو الأخرى بدون إرجاع . المطلوب تقدير (ا) احتمال أن الوحدتين غير تالفتين (ب) احتمال وجود وحدة تالفة .

– ٢٨ – صندوق يحتوى على 5 كرات سوداء و3 كرات خضراء . سحبت ثلاث كرات الواحد تلو الأخرى بدون إرجاع ما احتمال أن كل الكرات المسحوبة من نفس اللون ؟

– ٢٩ – إذا كان A , B حادثتين بحيث أن :-

$$P(A \cap B) = \frac{1}{10} \; , P(A^c \cap B) = \frac{1}{5}, P(A \cap B^c) = \frac{2}{3}$$

(ا) ما هو احتمال A^c , B^c ؟

(ب) هل الحادثتين مستقلتين ؟

– ٣٠ – مستحضر في أنبوبة اختبار يحتوى على عشرين من حبوب لقاح الصنوبر وخمسة من حبوب لقاح البلوط . اختيرت عينة عشوائية تحتوى على أربعة حبوب لقاح، ما هو احتمال أن :

(ا) تحتوى العينة على أربع حبوب من الصنوبر .

(ب) تحتوى العينة على ثلاث حبوب من البلوط .

(جـ) تحتوى العينة على الأقل على ثلاث حبوب لقاح الصنوبر .

– ٣١ – أطلق صياد 7 طلقات نارية على حيوان مفترس فإذا كان احتمال أن يصيب الهدف هو 0.6 ما هو احتمال أن الصياد ما زال على قيد الحياة ؟

– ٣٢ – صوب شخصان ناحية هدف ما، فإذا كان في المتوسط A يكسب 3 من 5 و B يكسب 4 من 8 ما هو احتمال أن الهدف لا يستهدف إذا صوب الاثنين ناحية الهدف ؟

– ٣٣ – تقدم ثلاثة أشخاص A , B , C إلى ثلاث وظائف مختلفة. احتمال أن يكسب A الوظيفة هو 0.8 واحتمال أن يكسب B الوظيفة هو 0.5 واحتمال أن يكسب C الوظيفة 0.45 . ما هو احتمال (ا) أن الثلاثة يحصلون على الوظائف ؟ (ب) عدم تعيين أى واحد في الوظيفة المقدم لها ؟ (جـ) واحد فقط يحصل على الوظيفة ؟

– ٣٤ – مجموعة مكونة من عشرة أشخاص، منهم ستة أشخاص مصابين بتسوس الأسنان. اختيرت منهم عينة عشوائية من ثلاثة أشخاص. ما هو احتمال أن تحتوى العينة على ثلاثة أشخاص مصابين بتسوس الأسنان.

– ٣٥ – آلة تتكون من ثلاثة أجزاء، الآلة تعتبر تالفة إذا كان واحد أو أكثر تالفا. احتمال أن الجزء A يتلف هو 0.01 واحتمال الجزء B يتلف هو 0.02 واحتمال الجزء C يتلف هو 0.1 أوجد احتمال (ا) الآلة تالفة (ب) أن تلف الآلة يرجع إلى فشل الجزء C فقط.

– ٣٦ – شركة طيران لها ست رحلات من بلد A إلى B وسبع رحلات من B إلى C (يوميا) ما عدد الرحلات التي تنجزها يوميا من A إلى C ؟

– ٣٧ – اختيرت ثلاثة فئران من مجموعة مكونة من خمسة فئران بيضاء اللون وأربعة بنية اللون لاستخدامها في تجربة معينة . ما هو احتمال أن تكون :–

(ا) جميع الفئران المختارة بيضاء اللون ؟

(ب) الفئران المختارة مكونة من فأر بنى وفأرين لونهما أبيض ؟

(جـ) جميع الفئران المختارة بنية اللون ؟

– ٣٨ – اختيرت ثلاث بذرات لنبات مزهر عشوائيا من كيس يحتوى على عشرة بذور زهورها حمراء وخمس زهورها بيضاء، ما هو احتمال أن تكون :–

(ا) زهور البذور الثلاثة من نفس اللون ؟

(ب) زهور البذور الثلاثة المختارة مختلفة الألوان ؟

– ٣٩ – في مستعمرة كبيرة لذبابة الفاكهة ، 20% من الذباب به طفرة في بالجناح، 35% به طفرة في العين، 10% به طفرة بكل من الجناح والعين. اختيرت ذبابة من المستعمرة عشوائيا. ما هو احتمال أن يكون بها أحد الطفرتين على الأقل ؟

– ٤٠ – حقنت ثمانية فئران بعقار معين وتم رصد عدد الفئران التي ماتت خلال يوم. إذا كان احتمال موت ستة بالضبط هو 0.03 واحتمال أن يموت سبعة أو ثمانية هو 0.04 . أوجد احتمال أن :–

(ا) يموت ستة فئران أو أكثر .

(ب) يموت خمسة أو أقل .

– ٤١ – إذا عُلم أن احتمال أن يكون الجو في بلد ما في شهر يناير ملبداً بالغيوم هو 0.6 واحتمال أن يكون الجو عاصفاً هو 0.65 واحتمال أن يكون ملبداً بالغيوم وعاصفاً هو 0.25 أوجد الاحتمالات الآتية :–

(ا) أن يكون الجو ملبدا بالغيوم وغيرعاصف .

(ب) أن يكون غير ملبد بالغيوم وغير عاصف .

(جـ) أن يكون الجو غير ملبد بالغيوم وعاصف .

– ٤٢ – بفرض أن 100 مستودع تم تصنيفهم حسب الإدارة (A,B,C) وحسب المبيعات (عالي، متوسط، منخفض) في الجدول المزدوج التالي :–

	الإدارة			
المبيعات	A	B	C	المجموع
عالى	20	4	2	26
متوسط	4	46	14	64
منخفض	1	2	7	10
المجموع	25	52	23	100

أستخدم البيانات الموجودة في الجدول في حساب :–

(ا) احتمال أن المبيعات عالية .

(ب) احتمال أن المبيعات متوسطة إذا علم أن الإدارة A .

(حـ) احتمال أن المبيعات متوسطة إذا علم أن الإدارة B .

– ٤٣ – في استطلاع للرأي عن تأثير الإعلانات على البيع في مركز لتسويق الأغذية، أُخذت عينة من 230 فرد من المترددين على المركز وسُجلت إجابتهم . الجدول المزدوج التالي يوضح توزيع الأفراد حسب الشراء (يشترى ولا يشترى) وحسب مشاهدة الإعلانات (يشاهد ولا يشاهد) . سحبت استمارة عشوائيا فإذا علم أن الشخص المختار يشاهد الإعلانات ما هو احتمال أن يشترى .

	يشترون	لا يشترون	المجموع
يشاهد	80	100	180
لايشاهد	20	30	50
المجموع	100	130	230

– ٤٤ – في بحث ميداني لدراسة العلاقة بين العمر و استخدام حزام الأمان في مدينة بها 1000 وقائد سيارة تم الحصول على البيانات التالية :–

	لا يستخدمون الحزام	يستخدمون الحزام	المجموع
تحت 40	250	177	427
40 فأكبر	325	248	573
المجموع	575	425	1000

(ا) ما هو احتمال أن قائد السيارة يستخدم حزام الأمان ؟

(ب) ما هو احتمال أن قائد السيارة تحت 40 سنة ولا يستخدم حزام الأمان ؟

(جـ) إذا علم أن قائد السيارة يستخدم حزام الأمان ما هو احتمال أن عمره 40 فاكبر.

– ٤٥ – في متجر لبيع الملابس النسائية تم تصنيف 230 فرد من المترددين على المتجر حسب الشراء إلى (يشترون ولا يشترون) وحسب الجنس (ذكور إناث)كما في الجدول المزدوج التالي :–

(ا) ما هو الاحتمال أن المشترى أنثى ؟

(ب) إذا تم الشراء ما هو احتمال أن المشترى أنثى ؟

(جـ) هل إمكانية الشراء مستقلة عن نوع المشترى (ذكر أو أنثى) ؟

	يشترون	لا يشترون
إناث	80	100
ذكور	20	30

– ٤٦ – طائرة تطير يوميا بين مدينتي فإذا كان احتمال أن تقوم في ميعادها 0.8 ، احتمال أن يكون الجو جيد عندما تطير في موعدها هو 0.9 ،وعندما لا تطير في موعدها فإن احتمال أن يكون الجو رديئا هو 0.7 ،إذا ركب شخص الطائرة وكان الجو جيد ما هو احتمال أن الطيران يكون في ميعاده ؟

– ٤٧ – بفرض أن 1% من سكان مدينة ما يعانون من مرض ما ، فإذا ظهر اختبار جديد للكشف عن المرض وأجرى على سكان المدينة، أعطى الاختبار نتيجة موجبة في 95% من الحالات التي عندها المرض، كما أعطى الاختبار نتيجة سالبة في 97% من الحالات التي ليس عندها المرض، اختير شخص بطريقة عشوائية وكانت نتيجة الكشف عنده موجبة، ما احتمال أنه يعانى من المرض.

٤٨- – مصنع ينتج ثلاثة أصناف من المصابيح بنسب 60% , 30% , 10% • التالف في الإنتاج هي 4% , 3% , 2% على التوالي اختير إحدى أصناف الإنتاج واختير منه مصباح أوجد:-

(ا) احتمال أن المصباح تالف •

(ب) إذا كان المصباح تالف أوجد احتمال أن يكون من إنتاج الصنف الأول •

– ٤٩ – تمثل الطالبات 30% من حجم الدارسين في كلية ما • يدرس 30% من الطلاب مادة الرياضيات وتدرس 20% من الطالبات مادة الرياضيات • إذا اختير واحد من الكلية ووجد أنه يدرس الرياضيات ما هو احتمال أن يكون المختار طالبة •

– ٥٠ – في دراسة ميدانية في إحدى الكليات وجد أن 7% من الذكور ، 2% من الإناث أطول من 1.7 مترا وأن 70% من الدارسين من الإناث • اختير واحد عشوائيا ووجد أنه أطول من 1.7 فما احتمال أن تكون أنثى •

– ٥١ – إذا كان 20% من العاملين في شركة ما يحملون شهادات عليا وإذا كان 25% من الذين يحملون شهادات متوسطة يشغلون مناصب عليا • أيضا 75% من الذين يحملون شهادات عالية يشغلون مناصب عليا • فإذا اختير أحد الأشخاص عشوائيا ما هو احتمال أن يحمل شهادة عليا إذا علم أنه يشغل منصب عالى •

– ٥٢ – لدينا ثلاث أوعية كما يلي :– الوعاء الأول به 4 كرات بيضاء، والوعاء الثاني به ثلاث كرات حمراء و 4 بيضاء ، والوعاء الثالث به 2 حمراء و 3 بيضاء اختير وعاء بطريقة عشوائية وسحبت منه كرة ووجد أن الكرة بيضاء فما احتمال أن تكون من الوعاء الأول ؟

– ٥٣ – يذهب رجل إلى عمله يوميا إما بسيارته أو بوسائل النقل العام • احتمال أن يركب سيارته هو 0.4 واحتمال أن يتأخر عن عمله إذا أستخدم وسائل النقل العام هو 0.3 واحتمال أن يتأخر عن عمله إذا أستخدم سيارته هو 0.1 فإذا ذهب إلى عمله متأخرا في يوما ما أوجد احتمال أن يكون قد أستقل سيارته •

الفصل الرابع

المتغيرات العشوائية وتوزيعاتها الاحتمالية

Random Variables and their Probability Distributions

(٤-١) المتغير العشوائي Random Variable

تستخدم كلمة تجربة (كما ذكرنا سابقا) لأي إجراء نعلم مسبقا جميع النواتج الممكنة له وإن كنا لا نستطيع أن نتنبأ بأي من هذه النواتج سيتحقق فعلا. ربما لا يكون من الضروري دراسة فئة كل النواتج الممكنة (فراغ العينة) لتجربة إحصائية ولكن يكون اهتمامنا منصبا على قيم رقمية مرتبطة بهذه النواتج الممكنة. إن القيم الممكنة هذه هي ما نعبر عنه بقيم المتغير العشوائي.

تعريف : الدالة المعرفة على فراغ العينة لتجربة ما والتي تخصص عددا حقيقيا لكل نقطة عينة تسمى المتغير العشوائي.

سوف نستخدم الرمز X ليمثل المتغير العشوائي، x لواحدة من قيمه.

مثال (٤-١) اختيرت بذرتان من نبات مزهر عشوائيا من كيس يحتوى على خمس بذور زهورها حمراء وثلاث بذور زهورها صفراء وذلك لاستخدامها في تجربة معينة. فراغ العينة يكون :

$$S = \{yy, ry, yr, rr\}$$

حيث r ترمز إلى البذرة التي زهورها حمراء، y ترمز إلى البذرة التي زهورها صفراء. بفرض أننا عرفنا الدالة X التي تمثل عدد البذور التي زهورها حمراء في العينة. هذه الدالة سوف تخصص عددا حقيقيا لكل نقطة عينة في فراغ العينة S المرافق لتجربتنا الإحصائية. في الجدول التالي نجد أن كل نقطة في فراغ العينة ارتبطت بعدد حقيقي واحد عن طريق الدالة X.

x	0	1	1	2
نقطة العينة	yy	ry	yr	rr

وعلى ذلك X متغير عشوائي يأخذ القيم $0 , 1 , 2$

قد يحتوى فراغ العينة على عدد محدود من النقط كما في المثال السابق، أو قد يكون فراغ العينة لانهائي معدود **countable infinite sample space** وهو الفراغ الذي يحتوى على عدد لانهائي من العناصر لكنه قابل للعد بمعنى أن هناك تقابل بين عناصره وفئة الأعداد الطبيعية ، مثل عدد البكتريا في لتر من الماء النقي أو عدد الفئران في فدان من القمح. يسمى فراغ العينة في هذه الحالة فراغ عينة منفصل (متقطع) **discrete sample space** . المتغير العشوائي المعرف على فراغ عينة منفصل يسمى متغير عشوائي منفصل (متقطع) **discrete random variable**. أيضا إذا كان فراغ العينة يحتوى على عدد لانهائي من النقط **infinite sample space** الغير معدودة مثل كل الأطوال الممكنة ، الأوزان ، درجات الحرارة ،

العمر ٠٠٠الخ فإننا نقول أن فراغ العينة متصل (مستمر) continuous sample space ٠ المتغير العشوائي المعرف على فراغ عينة متصل يسمى المتغير العشوائي المتصل continuous random variable. في معظم التطبيقات المتغيرات العشوائية المنفصلة تمثل بيانات قابلة للعد ، مثل عدد الحوادث في السنة ، عدد الأخطاء في صفحة من قاموس، عدد الفئران في فدان من القمح٠٠٠الخ٠ أما المتغيرات العشوائية المتصلة فتمثل بيانات مقاسة٠

(٤-٢) التوزيعات الاحتمالية المنفصلة (المتقطعة)

Discrete Probability Distributions

كل قيمة من قيم المتغير العشوائي المنفصل يفرض لها احتمال ففي مثال (٤ -١) تحسب الاحتمالات المختلفة لقيم المتغير العشوائي X الذي يمثل عدد البذور التي زهورها حمراء في العينة (إذا كان الاختيار بدون إرجاع) كالتالي :-

$$P(X=0)=P(yy)=(\frac{3}{8})(\frac{2}{7})=\frac{6}{56},$$
$$P(X=1)=P(ry)+P(yr)$$
$$=(\frac{5}{8})(\frac{3}{7})+(\frac{3}{8})(\frac{5}{7})=\frac{30}{56},$$
$$P(X=2)=P(rr)=(\frac{5}{8})(\frac{4}{7})=\frac{20}{56}.$$

القيم المختلفة للمتغير العشوائي X مع احتمالاتها معطاة في الجدول التالي :-

x	0	1	2
P(X=x)	$\frac{6}{56}$	$\frac{30}{56}$	$\frac{20}{56}$

مجموع الاحتمالات في الجدول السابق تساوى الواحد الصحيح٠

مثال (٤-٢) المطلوب الاحتمالات المختلفة لقيم المتغير العشوائي X في مثال (٤-١) إذا كان الاختيار بإرجاع :-

الحل٠

$$P(X=0)=P(yy)=(\frac{3}{8})(\frac{3}{8})=\frac{9}{64},$$
$$P(X=1)=P(ry)+P(yr)$$
$$=(\frac{5}{8})(\frac{3}{8})+(\frac{3}{8})(\frac{5}{8})=\frac{30}{64},$$
$$P(X=2)=P(rr)=(\frac{5}{8})(\frac{5}{8})=\frac{25}{64}.$$

القيم المختلفة للمتغير العشوائي X مع احتمالاتها معطاة في الجدول التالي :-

x	0	1	2
P(X=x)	$\frac{9}{64}$	$\frac{30}{64}$	$\frac{25}{64}$

عادة يفضل تمثيل كل احتمالات المتغير العشوائي X بصيغة. هذه الصيغة من الضروري أن تكون دالة في القيم الرقمية x . سوف نرمز للدالة بواحدة من الصيغ $f(x), g(x), h(x)$ ...الخ . وعلى ذلك يمكن كتابة $f(x) = P(X = x)$ ، فعلى سبيل المثال $f(2) = P(X = 2)$. الفئة من الأزواج المرتبة (x , f(x)) تسمى دالة الاحتمال **probability function** أو التوزيع الاحتمالي **probability distribution** للمتغير العشوائي X.

تعريف : كل جدول أو صيغة تعطى جميع القيم التي يأخذها متغير العشوائي منفصل، مع احتمال كل قيمة منها يسمى توزيع احتمالي منفصل.

مثال (٤-٣) إذا ألقيت عملة مرة واحدة وكان X متغير عشوائي حيث $x = 1$ إذا كانت النتيجة صورة و $x = 0$ إذا كانت النتيجة كتابة. في هذه الحالة يمكن وضع صيغة للدالة $f(x)$ الخاصة بالمتغير العشوائي X على الشكل :-

$$f(x) = \frac{1}{2} \;,\; x = 0 \;,\; 1.$$

مثال (٤-٤) أوجد صيغة التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل عدد الصور التي تظهر عند إلقاء خمس عملات مرة واحدة ؟

الحل. عدد النقط في فراغ العينة سوف يكون 2^5 . (الأحداث متساوية في إمكانية الحدوث) . المقام لجميع الاحتمالات سوف يكون 32 . لإيجاد عدد الطرق للحصول على **x** من الصور عند إلقاء **5** عملات مرة واحدة فإننا نحتاج لمعرفة العدد الكلى لنقاط العينة في التجربة والتي تعطى **x** صور و **n-x** كتابة وهذا يساوى عدد الطرق لتبديل **x** من العناصر منها **x** من نوع (صورة) و **n-x** من نوع آخر (كتابة) ، وهذا يحدث بطرق عددها $\binom{5}{x}$ حيث x تأخذ القيم 0,1,2,3,4,5 وعلى ذلك :-

$$f(x) = P(X = x) = \frac{\binom{5}{x}}{32} \;,\; x = 0,1,2,3,4,5.$$

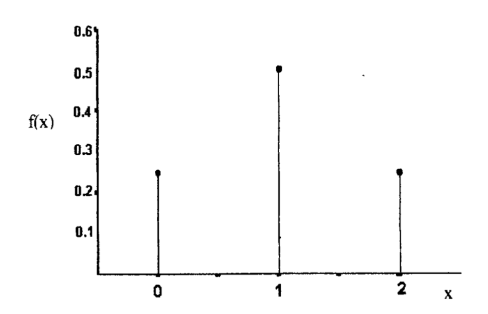
مثال (٤-٥) إذا كان X متغير عشوائي يمثل نواتج إلقاء زهرة نرد مرة واحدة فإن x تـأخذ قيم من 1 إلى 6 • التوزيع الاحتمالي للمتغير العشوائي X هو :-

$$f(x) = \frac{1}{6}\ ,\ x = 1,2,3,4,5,6.$$

مثال (٤-٦) إذا كان X متغيرا عشوائيا يمثل عدد الصورة التي تظهر عند إلقاء عملتـين مـرة واحدة فإن $x = 0,1,2$ • التوزيع الاحتمالي للمتغير العشوائي X يمكن تمثيله بالجدول التالي :

x	0	1	2
f(x)	0.25	0.5	0.25

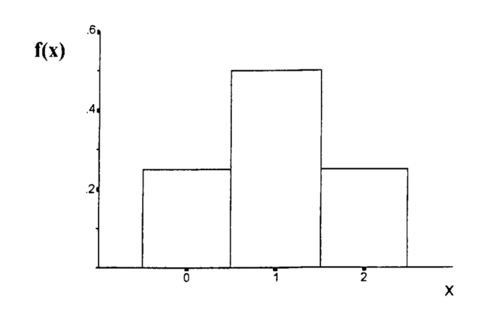
يمكن عرض هذا التوزيع بيانيا باستخدام طريقة الأعمدة **bar chart** كما في شكل (٤-١)•



شكل (٤-١)

حيث يمثل المحور الأفقي قيم x ويمثل المحور الرأسي قيم f(x) فمثلا عند $x = 0$ يقام عمـود إرتفاعة يتناسب مع قيمة الدالة عند هذه النقطة وهو 0.25 وكذلك عند $x = 1$ يقام عمـود ارتفاعه 0.5 وعند $x = 2$ يقام عمود ارتفاعه 0.25 وبخلاف هذه النقط فالدالة ليـس لهـا وجود• كما يمكن تحويل شكل (٤-١) إلى ما يسمى بالمدرج الاحتمـالي **probability histogram** كما في شكل (٤-٢) وذلك بتحويل الأعمدة الموجودة إلى مستطيلات بحيـث يكون ارتفاع كل مستطيل مساويا لاحتمال قيمة x الواقعة في منتصف قاعدة المستطيل• وعلـى

ذلك فإن $P(X = x)$ يساوى مساحة المستطيل الذي تقع x في منتصف قاعدته . هذا المفـــهوم لحساب الاحتمالات ضروري في التوزيع الاحتمالي المتصل .

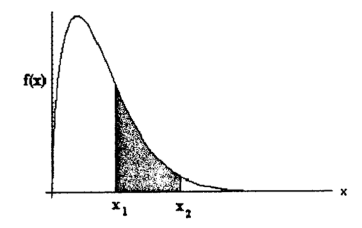


شكل (٤-٢)

(٤-٣) التوزيعات الاحتمالية المتصلة (المستمرة)

Continuous Probability Distributions

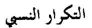
من صفات المتغير العشوائي المتصل أنه لا يمكن أن يكون هناك احتمال موجب مرافـــق لكل قيمة من قيم المتغير أي أن $P(X = x) = 0$ ولكن يكون هناك احتمال مرافق لكل فـترة من فترات المتغيز . ولهذا لا يمكن تمثيل التوزيع الاحتمالي للمتغير العشوائي المتصل بجدول ولكـــن نعبر عنه بصيغة دالة $f(x)$ والتي تسمى دالـــة كثافـــة الاحتمـــال **probability density function** . التمثيل البياني للدالة $f(x)$ سوف يكون متصل ويأخذ أشكال كثيرة . واحـــد من هذه الأشكال موضح في شكل (٤-٣) ولابد أن تكون المساحة تحت منحنى الدالـة $f(x)$ والمحددة بمحور x تساوى الواحد الصحيح . أيضا احتمال أن يأخذ المتغير العشوائي X قيمـــة بين $x = x_1$ و $x = x_2$ يساوى المساحة المظللة تحت دالة كثافـــة الاحتمـــال بـــين $x = x_1$ و $x = x_2$. المساحة بالضبط يمكن الحصول عليها باستخدام طرق التكامل . ســـوف يكـــون اهتمامنا فقط بدوال كثافة الاحتمال الشائعة الاستخدام في التجارب والتي تحسب المساحات تحت منحناها باستخدام الجداول الإحصائية . ولأن المساحات تمثل احتمـــالات والاحتمـــالات قيـــم موجبة، فإن دالة كثافة الاحتمال لابد أن تكون فوق منحنى x .

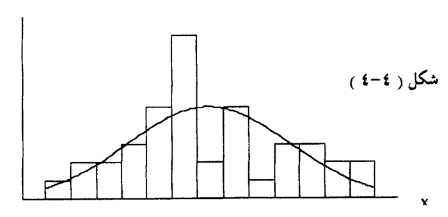


شكل (٤-٣)

تعريف : الدالة $f(x)$ تسمى دالة كثافة الاحتمال لمتغير عشوائي متصل X إذا كانت المساحة الكلية تحت المنحنى والمحددة بمحور x تساوى الواحد الصحيح. أيضا المساحة تحت المنحنى بين أي قيمتين $x = x_2$ و $x = x_1$ تعطى احتمال أن المتغير العشوائي يقع بين $x = x_1$ و $x = x_2$.

عادة في التجربة التي تحتوى على متغير متصل تكون $f(x)$ غير معلومة ويفرض معادلتها. الاختيار الصحيح للمعادلة يعتمد على توفر معلومات عن المتغير موضع الدراسة. المدرج يلقى الضوء على شكل $f(x)$. على سبيل المثال إذا كان الاهتمام بوزن مجموعة من الحيوانات فإننا نقوم بتوزيع المشاهدات (الأوزان x) إلى فئات ونقدر نسبة المشاهدات في كل فئة (فترة) والتي تسمى التكرارات النسبية **relative frequencies** للفئات المختلفة. هذه المعلومات يمكن تمثيلها بالمدرج ومنها نحصل على $f(x)$ التقريبية وذلك بتمهيد المنحنى كما في شكل (٤-٤). شكل المنحنى يساعدنا في الاختيار المناسب للدالة $f(x)$.





شكل (٤-٤)

(٤-٤) التوقع الرياضي Mathematical Expectation

يمكن تسهيل فهم التوقع الرياضي بالمثال التالي : ليكن X متغير عشوائي يمثل عدد الصور التي تظهر عند إلقاء عمليتين مرة واحدة • وعلى ذلك X يأخذ القيم 0,1,2 باحتمالات 0.25,0.5, 0.25 على التوالي • بفرض أن التجربة كررت بعدد كبير جدا من المرات، وليكن n = 8000000 ، نتوقع أن نلاحظ تقريبا 2 مليون للحادثة "عدم ظهور الصورة " و 4 مليون للحادثة "ظهور صورة و كتابة " و 2 مليون للحادثة "ظهور صورتين ". وعلى ذلك متوسط عدد الصور في الرمية الواحدة يساوى:

$$\frac{\text{Sum of observations}}{N} = \frac{(0)(2000000)+(1)(4000000)+(2)(2000000)}{8000000}$$

$$= \frac{(0)(2000000)}{8000000} + \frac{(1)(4000000)}{8000000} + \frac{(2)(2000000)}{8000000}$$

$$= (0)(\frac{1}{4}) + (1)(\frac{1}{2}) + (2)(\frac{1}{4}) = 1.$$

حيث عدد المشاهدات = sum of observations • يلاحظ أن الحد الأول يساوى $(0)f(0)$ والحد الثاني يساوى $(1)f(1)$ والحد الثالث يساوى $(2)f(2)$ وعلى ذلك يمكن تعريف القيمة المتوقعة للمتغير العشوائي X (متوسط التوزيع أو متوسط المجتمع) كالتالي :-

$$\mu = E(X) = (0)(.25) + (1)(0.5) + (2)(0.25) = 1.$$

تعريف : إذا كان X متغير عشوائي منفصل له التوزيع الاحتمالي التالي :-

x	x_1	x_2	...	x_n
P(X=x)	$f(x_1)$	$f(x_2)$	...	$f(x_n)$

فإن التوقع الرياضي (القيمة المتوقعة أو متوسط المجتمع μ population mean) لمتغير عشوائي X هو :-

$$E(X) = \sum_{i=1}^{n} x_i f(x_i).$$

مثال (٤-٧) اختيرت عينة من 3 وحدات بطريقة عشوائية من صندوق به 12 وحدة بينها 3 معيبة أوجد القيمة المتوقعة لعدد الوحدات المعيبه •

الحل ، يتكون فراغ العينة S من $220 = \binom{12}{3}$ عينة متساوية الاحتمال حجمها 3 ، أيضا يوجد $84 = \binom{9}{3}\binom{3}{0}$ عينة ليست بها وحدات معيبة و $108 = \binom{9}{2}\binom{3}{1}$ عينة بها وحدة واحدة معيبة و $27 = \binom{9}{1}\binom{3}{2}$ عينة بها وحدتين معيبتين وعينة واحدة، $1 = \binom{9}{0}\binom{3}{3}$، بها ثلاث وحدات معيبة وبذلك يكون احتمال الحصول على 3,2,1,0 من الوحدات المعيبة على التوالي هو :-

$\frac{84}{220}, \frac{108}{220}, \frac{27}{220}, \frac{1}{220}$ ، أذن القيمة المتوقعة لعدد الوحدات المعيبة هي :-

$$E(X) = (0)f(0) + (1)f(1) + (2)f(2) + (3)f(3)$$

$$= (0)(\frac{84}{220}) + (1)(\frac{108}{220}) + (2)(\frac{27}{220}) + (3)(\frac{1}{220}) = \frac{165}{220}.$$

مثال (٤-٨) يحتوى صندوق على 4 كرات ، اثنين منهم مرقمين بالرقم 2 ، وكرة مرقمة بـالرقم 4 ، والأخيرة مرقمة بالرقم 8 سحبت كرة من الصندوق أوجد $E(X)$.

الحل، قيم المتغير العشوائي سوف تكون $x = 2,4,8$ باحتمالات :-

$$f(2) = \frac{1}{2}, f(4) = \frac{1}{4}, f(8) = \frac{1}{4}.$$

وعلى ذلك القيمة المتوقعة للمتغير X أو متوسط المجتمع هو :-

$$\mu = E(X) = (2)(\frac{1}{2}) + (4)(\frac{1}{4}) + (8)(\frac{1}{4}) = 4.$$

في الجزء السابق تناولنا كيفية إيجاد القيمة المتوقعة لمتغير عشوائي، ليكـن $h(X)$ دالـة في متغير عشوائي (متغيرا عشوائيا جديد يعتمد على X)، وعلى ذلك يمكن تقدير قيم الدالة $h(X)$ بمعرفة قيم X، على سبيل المثال $h(X)$ قد تكون $(X-\mu)^2$ أو X^2 ، وعلى ذلك إذا كـان X يـأخذ القيمـة 4 فـإن $h(X)$ تـأخذ القيمـة $h(4)$ وعلـى ذلـك يمكـن كتابـة :-
$P[h(X) = h(4)] = P(X = 4)$ ، عمومـا $P[h(X) = h(x)] = P(X = x) = f(x)$.
وعلى ذلك يمكن الحصول على القيمة المتوقعة للدالة $h(X)$ باستخدام الاحتمالات المعطاة للدالـة f(x)، على سبيل المثال إذا كان X متغيرا عشوائيا يمثل عدد الصور التي تظهر عند إلقاء عملتين مرة واحدة وبفرض أننا نرغب في إيجاد القيمة المتوقعة للدالة $Y = X^2$ ، التوزيع الاحتمالي للمتغير X وأيضا للمتغير العشوائي Y في الجدول التالي :-

x	0	1	2
y	0	1	4
P(X=x) =P(Y=y)	0.25	0.5	0.25

$$E(Y) = (0)(\frac{1}{4}) + (1)(\frac{1}{2}) + (4)(\frac{1}{4})$$

$$= \sum_{x=0}^{2} x^2 f(x) = 1.5.$$

تعريف : إذا كان X متغير عشوائي منفصلا له التوزيع الاحتمالي التالي :-

x	x_1	x_2	...	x_n
P(X=x)	$f(x_1)$	$f(x_2)$	...	$f(x_n)$

وإذا كان $h(X)$ دالة في X فإن $h(X)$ تمثل أيضا متغيرا عشوائيا والقيمة المتوقعة له هي:-

$$E[h(X)] = \sum_{i=1}^{n} h(x_i) f(x_i)$$

مثال (٤-٩) إذا كان X متغيرا عشوائيا بتوزيع احتمالي معطى في الجدول التالي أوجد القيمة المتوقعة للمتغير $(X^2 - 1)$:-

الحل.

x	-1	0	1	2
P(X=x)	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$

$$E(X^2 - 1) = \sum_{x=-1}^{2} (x^2 - 1) f(x)$$

$$= [(-1)^2 - 1]f(-1) + [(0)^2 - 1]f(0) + [(1)^2 - 1]f(1) + [(2)^2 - 1]f(2)$$

$$= (0)(\frac{1}{8}) + (-1)(\frac{1}{4}) + (0)(\frac{3}{8}) + (3)(\frac{1}{4}) = \frac{1}{2}.$$

مثال (٤-١٠) أوجد القيمة المتوقعة للدالة $(X - \mu)^2$ حيث أن X تمثل عدد الصور التي تظهر عند إلقاء عملتين مرة واحدة.

الحل. قد أثبتنا من قبل أن $E(X) = 1$ ، وعلى ذلك فإن :-

$$E(X-\mu)^2 = \sum_{x=0}^{2} (x \ -\mu)^2 f(x)$$

$$= (0-1)^2 f(0) + (1-1)^2 f(1) + (2-1)^2 f(2)$$

$$= (1)(0.25) + (0)(0.5) + (1)(0.25) = 0.5.$$

القيمة المتوقعة للدالة $(X-\mu)^2$ تسمى التباين **variance** للمتغير العشوائي X ويرمز لها بالرمز σ^2 ، يعرف الانحراف المعياري **standard deviation** للمتغير العشوائي X بأنه الجذر التربيعي لتباين X .

تعريف : التباين للمتغير العشوائي X هو :-

$$\sigma^2 = E(X-\mu)^2.$$

مثال (٤-١١) الجدول التالي يمثل التوزيع الاحتمالي للمتغير العشوائي X والذي يمثل عدد أجهزة الحاسب الآلي من 4 أجهزة والتي قد تتعرض للتلف أثناء عملية الشحن إلى مركز أبحاث ، أوجد التباين والانحراف المعياري للمتغير العشوائي X .

x	0	1	2	3	4
P(X=x)	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

الحل ، بما أن التباين للمتغير العشوائي X معرف بالصيغة $\sigma^2 = E(X-\mu)^2$ ، أولا نحسب العدد المتوقع للأجهزة التالفة كالتالي :-

$$\mu = E(X) = \sum_{x=0}^{4} x\, f(x)$$

$$= 0\cdot(\frac{1}{16}) + 1\cdot(\frac{4}{16}) + 2\cdot(\frac{6}{16}) + 3\cdot(\frac{4}{16}) + 4\cdot(\frac{1}{16})$$

$$= \frac{4}{16} + \frac{12}{16} + \frac{12}{16} + \frac{4}{16} = \frac{32}{16} = 2.$$

الآن :-

$$\sigma^2 = E(X-\mu)^2 = \sum_{x=0}^{4} (x \quad -2)^2 f(x \quad)$$

$$= (0-2)^2(\frac{1}{16}) + (1-2)^2(\frac{4}{16}) + (2-2)^2(\frac{6}{16}) + (3-2)^2(\frac{4}{16}) + (4-2)^2(\frac{1}{16})$$

$$= (4)(\frac{1}{16}) + (1)(\frac{4}{16}) + (0)(\frac{6}{16}) + (1)(\frac{4}{16}) + (4)(\frac{1}{16}) = 1.$$

أيضا الانحراف المعياري للمتغير X هو $\sigma = \sqrt{\sigma^2} = 1$ ٠

تعريف : العزم من الدرجة k (حيث k عدد صحيح موجب) حول نقطة الأصل للمتغير العشوائي X هو :-

$$\mu'_k = E(X^k).$$

والعزم من الدرجة k حول المتوسط هو :-

$$\mu_k = E(X-\mu)^k.$$

العزم من الدرجة الأولى حول الصفر يعطى متوسط المجتمع μ والعزم من الدرجة الثانية حول المتوسط يعطى التباين σ^2 ٠ العزوم بصفة عامة لهم استخدامات كثيرة في الإحصاء سوف نتناول بعضها في الفصل الخامس٠

(٤-٥) بعض خواص القيم المتوقعة

Some Properties of Expected Values

في هذا البند سوف نقدم بعض النظريات والتي عن طريقها يمكن حساب توقعات بدلالة توقعات أخرى معروفة أو توقعات سهلة في الحساب٠ كل النتائج التالية صحيحة سواء لمتغيرات عشوائية منفصلة أو متصلة٠ البراهين التالية سوف تقتصر على المتغيرات العشوائية المنفصلة المحدودة٠

نظرية (٤-١) بفرض أن X متغيرا عشوائيا و b, a ثابتين فإن :-

$$E(aX+b) = aE(X) + b.$$

البرهان :-

$$E(aX+b)=\sum_{i=1}^{n}(ax_i+b)f(x_i)$$
$$=(ax_1+b)f(x_1)+(ax_2+b)f(x_2)+\ldots+(ax_n+b)f(x_n)$$
$$=a[x_1f(x_1)+x_2f(x_2)+\ldots+x_nf(x_n)]$$
$$+b[f(x_1)+f(x_2)+\ldots+f(x_n)]$$
$$=a\sum_{i=1}^{n}x_if(x_i)+b\sum_{i=1}^{n}f(x_i).$$

المجموع الأول من اليمين هو $E(X)$ والمجموع الثاني يساوى واحد صحيح • وعلى ذلك فإن :-

$$E(aX+b)=aE(X)+b.$$

نتيجة (١) إذا كانت $a=0$ ، فإن $E(b)=b$ •

نتيجة (٢) إذا كانت $b=0$ ، فإن $E(aX)=aE(X)$ •

نظرية (٤-٢) التوقع الرياضي لمجموع دالتين (أو أكثر) في متغير عشوائي X تساوى مجموع القيم المتوقعة للدوال • أي أن :-

$$E[g(X)+h(X)]=E[g(X)]+E[h(X)].$$

البرهان :-

$$E[g(X)+h(X)]=\sum_{i=1}^{n}[g(x_i)+h(x_i)]f(x_i)$$
$$=[g(x_1)+h(x_1)]f(x_1)+[g(x_2)+h(x_2)]f(x_2)+\ldots$$
$$+[g(x_n)+h(x_n)]f(x_n)$$
$$=[g(x_1)f(x_1)+g(x_2)f(x_2)+\ldots+g(x_n)f(x_n)]$$
$$+[h(x_1)f(x_1)+h(x_2)f(x_2)+\ldots+h(x_n)f(x_n)]$$
$$=\sum_{i=1}^{n}g(x_i)f(x_i)+\sum_{i=1}^{n}h(x_i)f(x_i)$$
$$=E[g(X)]+E[h(X)].$$

نظرية (٤-٣) التباين للمتغير العشوائي X يعطى من الصيغة التالية :-

$$\sigma^2=E(X)^2-\mu^2.$$

البرهان :-

$$\sigma^2 = E(X-\mu)^2$$
$$= E(X^2 - 2\mu x + \mu^2)$$
$$= E(X^2) - 2\mu E(X) + E(\mu^2)$$
$$= E(X)^2 - \mu^2.$$

حيث أن $E(X) = \mu$ من التعريف و $E(\mu^2) = \mu^2$ من نظرية (٤-١) ونتيجة (١) .

مثال (٤-١٢) أوجد التباين للمتغير العشوائي X في مثال (٤-١٠).

الحل . قد أثبتنا من قبل أن $E(X) = 1$ ، الآن :-

$$E(X^2) = \sum_{x=0}^{2} x^2 f(x) = (0)(0.25) + (1)(0.5) + (2)^2(0.25) = 1.50.$$

وعلى ذلك :-

$$E(X-\mu)^2 = E(X^2) - \mu^2$$
$$= (1.5) - (1)^2 = 0.5.$$

وهي نفس النتيجة التي حصلنا عليها من مثال (٤-١٠).

مثال (٤-١٣) أوجد التباين للمتغير العشوائي X في مثال (٤-١١) باستخدام الصيغة :-

$$E(X^2) - \mu^2.$$

الحل . $E(X) = 2$ تم الحصول عليه من مثال (٤-١١) الآن نوجد :-

$$E(X^2) = \sum_{x=0}^{4} x^2 f(x)$$

$$= (0)^2(\frac{1}{16}) + (1)^2(\frac{4}{16}) + (2)^2(\frac{6}{16}) + (3)^2(\frac{4}{16}) + (4)^2(\frac{1}{16})$$
$$= \frac{4}{16} + \frac{24}{16} + \frac{36}{16} + \frac{16}{16} = \frac{80}{16} = 5.$$

وعلى ذلك فإن التباين للمتغير العشوائي X هو :-

$$\sigma^2 = E(X)^2 - \mu^2$$
$$= 5 - 2^2 = 1.$$

وهي نفس النتيجة التي حصلنا عليها من مثال (٤-١١).

تعريف : ليكن X متغير عشوائي بدالة كثافة احتمال $f(x)$. التباين لمتغير عشوائي جديد $g(X)$ هو:-

$$\sigma^2_{g(x)} = E[\{g(X) - \mu_{g(X)}\}^2].$$

نظرية (٤-٤) إذا كان X متغيرا عشوائيا و b ثابت ، فإن :-

$$\sigma^2_{X+b} = \sigma^2_X = \sigma^2.$$

البرهان :-

$$\sigma^2_{X+b} = E[\{(X+b) - \mu_{X+b}\}^2].$$

الآن

$$\mu_{X+b} = E(X+b) = E(X) + b$$
$$= \mu + b,$$

وذلك من نظرية (٤-١) وعلى ذلك :-

$$\sigma^2_{X+b} = E[(X + b - \mu - b)^2]$$
$$= E[(X - \mu)^2]$$
$$= \sigma^2.$$

نظرية (٤-٥) إذا كان X متغير عشوائي و a ثابت، فإن :-

$$\sigma^2_{aX} = a^2\sigma^2_X = a^2\sigma^2.$$

البرهان :-

$$\sigma^2_{aX} = E[\{aX - \mu_{aX}\}^2].$$

الآن :-

$$\mu_{aX} = E(aX) = aE(X) = a\mu,$$

وذلك من نظرية (٤-١) ، نتيجة (٢) وعلى ذلك :-

$$\sigma^2_{aX} = E[(aX - a\mu)^2]$$
$$= a^2E[(X - \mu)^2]$$
$$= a^2\sigma^2.$$

(٤-٦) التوزيعات الاحتمالية الثنائية المنفصلة

Discrete Bivariate Distributions

بفرض أن لدينا متغيرين عشوائيين Y, X بتوزيع احتمالي $h(y), g(x)$ على التوالي ٠ التوزيع الاحتمالي لوقوع Y, X في آن واحد عبارة عن صيغة دالة عادة يشار إليها بالرمز $f(x, y)$ وتسمى التوزيع الاحتمالي المشترك للمتغيرين Y, X ٠ وعلى ذلك في حالة التوزيع المنفصل.فإن٠ $f(x, y) = P(X = x, Y = y)$ ٠ أي أن $f(x, y)$ تعطي احتمال وقوع Y, X في آن واحد على سبيل المثال إذا ألقينا زهرتي نرد مرة واحدة وإذا كانت X تمثل النقط التي تظهر على السطح العلوي الزهرة الأولى و Y تمثل عدد النقط التي تظهر على السطح العلوي للزهرة الثانية٠ فالتوزيع الاحتمالي المشترك للمتغيرين X , Y هو :-

$$P(X = x , Y = y) = f(x,y) = \frac{1}{36} \quad , \quad x = 1,2,3,4,5,6, \quad y = 1,2,3,4,5,6.$$

تعريف : إذا كان Y, X متغيرين عشوائيين لهما دالة التوزيع الاحتمالي المشترك $f(x, y)$ ، فإن هذه الدالة تحقق الشروط التالية :-

(أ) $f(x, y) \geq 0$ لجميع القيم (x, y) ٠

(ب) $\sum\limits_{y}\sum\limits_{x} f(x, y) = 1$

من المثال السابق نجد أن $f(x, y) \geq 0$ لجميع القيم (x, y) حيث $f(x, y) = \frac{1}{36}$ لجميع القيم (x, y) ، $x = 1,2,3,4,5,6 \; ; \; y = 1,2,3,4,5,6$ و $\sum\limits_{y}\sum\limits_{x} f(x, y) = \sum\limits_{y}\sum\limits_{x} \frac{1}{36} = 1$

إذا كان Y, X متغيرين عشوائيين منفصلين لهما دالة التوزيع الاحتمالي المشترك $f(x, y)$ فإنه يمكن الحصول على التوزيع الاحتمالي للمتغير X على حدة وتوزيع Y على حدة ويسمى التوزيع في هذه الحالة بالتوزيع الهامشي٠ وعلى ذلك تكون دالة التوزيع الاحتمالي الهامشي للمتغير X هي :

$$g(x) = \sum_{y} f(x, y).$$

وبالمثل دالة التوزيع الاحتمالي الهامشي للمتغير Y هي :-

$$h(y) = \sum_{x} f(x, y).$$

مثال (٤-١٤) اختيرت عينة من شخصين لإجراء اختبار معين عليهم من بين مجموعة مكونة من أربعة غير مدخنين واثنين مدخنين. إذا كان X و Y معرفتان كالتالي x=0 إذا كان الشخص الأول غير مدخن و x=1 إذا كان الشخص الأول مدخنا. أيضا y=0 إذا كان الشخص الثاني غير مدخن و y=1 إذا كان الشخص الثاني مدخنا، فإن التوزيع الاحتمالي المشترك للمتغيرين X و Y في الجدول التالى. المطلوب إيجاد التوزيع الهامشي لكل من X و Y.

y \ x	0	1	h(y)
0	$\frac{6}{15}$	$\frac{4}{15}$	$\frac{10}{15}$
1	$\frac{4}{15}$	$\frac{1}{15}$	$\frac{5}{15}$
g(x)	$\frac{10}{15}$	$\frac{5}{15}$	

الحل. أولا، لحساب دالة التوزيع الهامشي للمتغير X نجمع الأعمدة كما يلي :-

$$g(0) = \sum_{y=0}^{1} f(0,y) = f(0,0) + f(0,1) = \frac{10}{15},$$

$$g(1) = \sum_{y=0}^{1} f(1,y) = f(1,0) + f(1,1) = \frac{5}{15}.$$

ثانيا، لحساب دالة التوزيع الهامشي للمتغير Y نجمع الصفوف كما يلي : -

$$h(0) = \sum_{x=0}^{1} f(x,0) = f(0,0) + f(1,0) = \frac{10}{15},$$

$$h(1) = \sum_{x=0}^{1} f(x,1) = f(0,1) + f(1,1) = \frac{5}{15}.$$

تعريف : إذا كان Y, X متغيرين عشوائيين بدالة كثافة احتمالية مشتركة، $f(x,y)$ فإن الدالة الاحتمالية المشروطة للمتغير Y بشرط أن $X = x$ تعرف بالصيغة :-

$$f_{Y|X}(y|x) = \frac{f(x,y)}{g(x)}.$$

لقيم x بحيث أن $g(x) > 0$.

وبنفس الشكل فإن الدالة الإحتمالية المشروطة للمتغير X بشرط أن $Y = y$ تعرف بالصيغة :-

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{h(y)}.$$

لقيم y بحيث أن $h(y) > 0$ •

مثال (٤-١٥) أوجد $f_{X|Y}(x \mid 1), f_{Y|X}(y \mid 0)$ للبيانات في مثال (٤-١٤) :-

الحل • أولا، الدالة $f_{X|Y}(x \mid 1)$ يمكن إيجادها كالتالي :-

$$f_{X|Y}(0 \mid 1) = \frac{f(0,1)}{h(1)} = \frac{4}{5},$$

$$f_{X|Y}(1 \mid 1) = \frac{f(1,1)}{h(1)} = \frac{1}{5}.$$

ثانيا، الدالة $f_{Y|X}(y \mid 0)$ يمكن إيجادها كالتالي :-

$$f_{Y|X}(0 \mid 0) = \frac{f(0,0)}{g(0)} = \frac{6}{10},$$

$$f_{Y|X}(1 \mid 0) = \frac{f(0,1)}{g(0)} = \frac{4}{10}.$$

تعريف : يكون المتغيرين العشوائيين **X** و **Y** مستقلين إذا كان :-

$$f(x,y) = g(x)h(y)$$

لكل قيم (x,y) •

مثال (٤-١٦) في المثال (٤-١٤) هل Y, X مستقلين ؟

الحل • Y, X غير مستقلين لأنه بالنظر إلى الجدول المعروض في مثال (٤-١٤) نجد :-

$$f(x,y) \neq g(x)h(y), x = 0,1, y = 0,1.$$

فعلى سبيل المثال $f(0,1) \neq g(0)h(1)$ حيث $f(0,1) = \frac{4}{15}, g(0) = \frac{10}{15}, h(1) = \frac{5}{15}$

نظرية (٤-٦) إذا كان Y, X متغيرين عشوائيين فإن :-

$$E(X \pm Y) = E(X) \pm E(Y).$$

فعلى سبيل المثال إذا ألقيت زهرة نرد مرتين وكانت X تمثل عدد النقط الذي تظهر على السطح العلوي في المرة الأولى و Y عدد النقط التي تظهر على السطح العلوي في المرة الثانية فإن $Y + X$ يمثل مجموع العددين اللذان يظهران على السطح العلوي للنرد عند إلقائها مرتين •

نظرية (٤-٧) إذا كان Y, X متغيرين عشوائيين مستقلين فإن :-

$$E(XY) = E(X)E(Y).$$

على سبيل المثال عند إلقاء نردين مرة واحدة فإن XY تمثل حاصل الضرب للعددين الظاهرين على النردين .

نظرية (٤-٨) بفرض أن المتغيرين العشوائيين X, Y مستقلين , فإن :-

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y.$$

البرهان :-

$$\sigma^2_{X+Y} = E[\{(X+Y) - \mu_{X+Y}\}^2].$$

الآن :-

$$\mu_{X+Y} = E(X+Y) = E(X) + E(Y) = \mu_X + \mu_Y,$$

وذلك من نظرية (٤-٦) وعلى ذلك :-

$$\begin{aligned}\sigma^2_{X+Y} &= E[\{(X+Y) - (\mu_X + \mu_Y)\}^2] \\ &= E[\{(X - \mu_X) + (Y - \mu_Y)\}^2] \\ &= E[(X - \mu_X)^2] + E[(Y - \mu_Y)^2] \\ &\quad + 2E[(X - \mu_X)(Y - \mu_Y)].\end{aligned}$$

الحدين الأولين يمثلان على التوالي σ^2_Y, σ^2_X . المطلوب إثبات أن الحد الأخير يساوى صفر .

وعلى ذلك :-

$$\begin{aligned}E[(X - \mu_X)(Y - \mu_Y)] &= E(XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y) \\ &= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X \mu_Y \\ &= E(XY) - \mu_X \mu_Y = 0,\end{aligned}$$

وذلك لان $E(XY) = E(X)E(Y)$ للمتغيرات المستقلة . وعلى ذلك:-

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y.$$

نتيجة (٣) بفرض أن X, Y متغيرين عشوائيين مستقلين، فإن :-

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y.$$

النتيجة نحصل عليها بوضع $X - Y$ على الشكل $X + (-Y)$ وعلى ذلك:-

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_{(-Y)}.$$

من نظرية (٤-٥) نعلم أن $\sigma^2_{(-Y)} = (-1)^2 \sigma^2_Y = \sigma^2_Y$ وعلى ذلك :-

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y.$$

تعريف : القيمة المتوقعة للدالة $(X-\mu_X)(Y-\mu_Y)$ تعرف بالتغاير بين المتغيرين Y, X ويرمز لها بالرمز $Cov(X,Y)$ أي أن :-

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

وهو يقيس درجة الترافق بين التمغيرين. بعض خواص التغاير معطاة في النظريات التالية.

نظرية (٤-٩) إذا كانت Y, X متغيرين عشواليين و b, a ثابتين فإن :-

$$Cov(aX, bY) = abCov(X,Y),$$
$$Cov(X+a, Y+b) = Cov(X,Y),$$
$$Cov(X, aX+b) = a\sigma^2_X.$$

مثال (٤-١٧) إذا كان Y, X متغيرين عشوائيين مستقلين حيث :-

$$\sigma^2_Y = 16, \sigma^2_X = 4, E(Y) = 3, E(X) = 2$$

أوجد : (أ) $E(5X-Y)$ (ب) σ^2_{X-Y} ، (جـ) $Cov(3X+2\ ,Y)$

(د) $Cov(X\ ,\ 5X-2)$

الحل. (أ) $E(5X-Y) = 5E(X) - E(Y) = (5)(2) - 3 = 7$

(ب)وحيث Y, X مستقلين فإن :-

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y$$
$$= 4 + 16 = 20.$$

(جـ) $Cov(3X+2\ , Y) = 3Cov(X,Y) = (3)(0) = 0$

(د) $Cov(X\ , 5X-2) = 5Cov(X,X) = 5\sigma^2_X = (5)(4) = 20$

نظرية (٤-١٠) إذا كان Y, X متغيرين عشوائيين، فإن :-

$$Cov(X,Y) = E(XY) - E(X)E(Y).$$

كما أن $Cov(X,Y) = 0$ إذا كان Y, X مستقلين بينما العكس ، عموما ، غير صحيح بمعنى أنه بالإمكان أن يكون $Cov(X,Y) = 0$ ولكن Y, X غير مستقلين .

نظرية (٤-١١) إذا كان Y, X متغيرين عشوائيين بدالة كثافة احتمال مشتركة $f(x,y)$ ، فإن:

$$\sigma^2_{X\pm Y} = \sigma^2_X + \sigma^2_Y \pm 2Cov(X,Y)$$

و $\sigma^2_{X\pm Y} = \sigma^2_X + \sigma^2_Y$ إذا كان Y, X مستقلين .

مثال (٤-١٨) إذا كان Y, X متغيرين عشوائيين بدالة كثافة احتمال :

$$f(x,y)=\frac{1}{4}$$

حيث $(x,y)=(0,1),(1,0),(0,-1),(-1,0)$ و $g(\pm1)=\frac{1}{4}, g(0)=\frac{1}{2}$ ودالة كثافة الاحتمال للمتغير Y نفس دالة كثافة الاحتمال للمتغير X وبما أن $E(XY)=0$ ، فإن :

$$Cov(X,Y)=E(XY)-E(X)E(Y)=0.$$

ولكن $f(1,0)\neq g(1)h(0)$، وعلى ذلك Y,X متغيرين غير مستقلين. عموما يمكن القول أن Y,X غير مستقلين إذا كان $Cov(X,Y)\neq 0$، ولكن $Cov(X,Y)=0$ لا يعنى أن المتغيرين Y,X مستقلين .

تعريف : إذا كان Y,X متغيرين عشوائيين بتباينى σ_X^2,σ_Y^2 وتغاير $Cov(X,Y)$ ، فإن معامل الارتباط بين Y,X هو :-

$$\rho=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}.$$

يقال للمتغيرين العشوائيين Y,X انهم غير مرتبطين إذا كان $\rho=0$ ، غير ذلك يقال أنهما مرتبطين.

نظرية (٤-١٢) إذا كان ρ معامل الارتباط بين المتغيرين Y,X فإن :-

$$-1\leq\rho\leq 1.$$

مثال (٤-١٩) إذا كان Y,X متغيرين عشوائيين بدالة كثافة احتمال مشتركة :-

$$f(x,y)=\frac{4}{5xy}, x=1,2 \text{ and } y=2,3.$$

أوجد : (ا) معامل الارتباط بين Y,X (ب) هل Y,X مستقلين أم لا ؟ .

الحل. (ا) التوزيع الاحتمالي المشترك للمتغيرين Y,X في الجدول التالي :

y \ x	1	2
2	$\frac{12}{30}$	$\frac{6}{30}$
3	$\frac{8}{30}$	$\frac{4}{30}$

وعلى ذلك :-

$$E(X)=(1)(\frac{20}{30})+(2)(\frac{10}{30})=\frac{4}{3},$$

$$E(Y)=(2)(\frac{18}{30})+(3)(\frac{12}{30})=\frac{12}{5},$$

$$E(X^2)=(1)^2(\frac{20}{30})+(2)^2(\frac{10}{30})=2,$$

$$E(Y^2)=(2)^2(\frac{18}{30})+(3)^2(\frac{12}{30})=6,$$

$$\sigma_X^2=E(X^2)-\mu_X^2=2-(\frac{4}{3})^2=\frac{2}{9},$$

$$\sigma_Y^2=E(Y^2)-\mu_Y^2=6-(\frac{12}{5})^2=\frac{6}{25},$$

$$E(XY)=(1)(2)(\frac{12}{30})+(2)(2)(\frac{6}{30})+(1)(3)(\frac{8}{30})+(2)(3)(\frac{4}{30})=\frac{48}{15},$$

وعلى ذلك :-

$$Cov(X,Y)=E(XY)-E(X)E(Y)=\frac{48}{15}-(\frac{4}{3})(\frac{12}{5})=0.$$

الارتباط بين Y, X هو :-

$$\rho=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}=\frac{0}{\sqrt{(\frac{2}{9})(\frac{6}{25})}}=0.$$

(ب) Y, X غير مستقلين لان $f(x,y)\neq g(x)h(y)$ لجميع قيم (x,y) ٠

تمارين

- ١ – صنف المتغيرات العشوائية التالية إلى منفصلة ومتصلة :-

(ا) الزمن اللازم لوصول طائرة (ب) الزمن اللازم لإنهاء امتحان (جـ) عدد المصابيح التالفة في صندوق يحتوى على 5 مصابيح (د) عدد الأخطاء التي يتعرض لها شخص ما عند كتابة خطاب على الآلة الكاتبة (ز) كمية اللبن الحليب التي تدرها بقرة في العام (ر) عدد البيض الذي تضعه دجاجة في الشهر ٠

– ٢ – ألقيت عملة متحيزة ثلاث مرات بحيث أن فرصة ظهور الصورة ضعف فرصة ظهور الكتابة. أوجد التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل عدد الصور التي تظهر على السطح العلوي للعملة .

– ٣ – الجدول التالي يمثل التوزيع الاحتمالي لمتغير عشوائي X :–

x	0	1	2
P(X=x)		0.4	0.2

(ا) ما هي قيمة $P(X=0)$ ؟ (ب) أوجد القيمة المتوقعة للمتغير X (جـ) $P(X>1)$.

– ٤ – إذا ألقيت زهرتي نرد مرة واحدة أوجد :–

(ا) التوزيع الاحتمالي للمتغير العشوائي Y الذي يمثل مجموع النقط التي تظهر على السطح العلوي للنردين ومثله بيانيا .

(ب) التوقع والتباين للمتغير Y .

(جـ) التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل القيمة المطلقة للفرق بين مجموع النقط التي تظهر على السطح العلوي النردين.

(د) التوقع والتباين للمتغير X .

– ٥ – إذا كان التوزيع الاحتمالي للزيادة في سعر سلعة ما في خلال سنة قادمة محددة كما في الجدول التالي حيث X=0 تعني عدم وجود زيادة و X=1 زيادة أقل من 3% و X=2 زيادة من 3% إلى 6% و X=3 زيادة أكثر من 6%.

x	0	1	2	3
P(X=x)	0.1	0.1	0.5	0.3

أوجد : التوقع والتباين للمتغير X .

– ٦ – أوجد الصيغة الاحتمالية للمتغير X الذي يمثل عدد الصور التي تظهر على السطح العلوي عند إلقاء عملة متزنة سبع مرات وأيضا التوقع والتباين للمتغير X .

– ٧ – بفرض أن شركة شحن اشترت سيارة كبيرة بمبلغ 15000 دولار. إذا فقدت السيارة سواء بالسرقة أو بحادثة فإن ذلك يمثل فقد كلى .الفرصة في الفقد 0.002 ، أوجد القيمة التوقعة للفقد (المتغير العشوائي هنا يأخذ القيمة 0 لعدم الفقد والقيمة 15000 للفقد).

– ٨ – أوجد القيمة المتوقعة لعدد الرجال الذين يتم اختيارهم لمهمة علمية من 3 أشخاص من بين 5 رجال وسيدتين.

– ٩ – إذا كان **X** متغيرا عشوائيا يمثل عدد أجهزة الحاسب الآلي التي تتعرض التلف من بين خمسة أجهزة وذلك أثناء توصيلها إلى مركز أبحاث. بفرض أن احتمال التلف **0.25**. أيضا بفرض أن كل جهاز مستقل عن الآخر في التلف أو عدم التلف أوجد القيمة المتوقعة للأجهزة التالفة.

– ١٠ – إذا كانت المبيعات من سلعة ما في الساعة هي **20 , 21 , 22** عبوة باحتمال **0.2 , 0.5 , 0.3** على التوالى. أوجد القيمة المتوقعة والتباين لعدد العبوات المباعة في الساعة.

– ١١ – احتمال أن يحصل لاعب كرة التنس على هدف في أى مباراة يلعبها هو **0.3**. أوجد القيمة المتوقعة لعدد الأهداف التي يكسبها في خمس مباريات قادمة.

– ١٢ – بفرض أن إيرادات متجر في اليوم تمثل بالمتغير العشوائي **X** ، والذي له التوزيع الاحتمالي التالي :–

x	0	10	12	16	18
P(X=x)	0.1	0.2	0.2	0.3	0.2

أوجد التوقع والتباين.

– ١٣ – إذا كانت دالة كثافة الاحتمال لمتغير عشوائي X هي :–

$$f(x) = \frac{1}{5}, x = 1,2,3,4,5.$$

أوجد احتمال أن X : (ا) عدد زوجي – عدد فردى (ب) التوقع والتباين للمتغير X .

– ١٤ – يقوم بائع بتوصيل نوعين من المنظفات (**B , A**) إلى المنزل. المكسب من المنظف **A** و**B** على التوالي هو **5 , 10** جنيهات للعبوة. الفرصة لبيع المنظف **A** هي **2** من **10** جولات والفرصة لبيع المنظف **B** هي **3** من **10** جولات والفرصة لعدم البيع هي **5** من **10** جولات. أوجد القيمة المتوقعة للمكسب في الجولة الواحدة.

– ١٥ – لدى محل للرياضة **80** علبة تحتوى كل علبة على كرات تنس ذات لون واحد ، إما صفراء أو خضراء. إذا كان عدد العلب التي تحتوى على كرات صفراء **30** . سحبت عينة من **10** علب. أوجد : (ا) التوزيع الاحتمالي لعدد العلب التي تحتوى على كرات صفراء (ب) القيمة المتوقعة لعدد العلب التي تحتوى على كرات صفراء.

– ١٦ – أى من الدوال التالية تمثل توزيع احتمالي :–

$$f(x) = \frac{x+2}{15} \;,\; x = -2,-1,0,1,2 \;\text{(ب)}\quad f(x) = \frac{2x}{5} \;,\; x = 0,1,2 \;\text{(ا)}$$

جـ) $f(x)=\frac{x^2-1}{49}\ ,\ x=0,1,2,3,4,5$ د) $f(x)=\frac{x}{3}\ ,\ x=-1,0,1,2$

ز) $f(x)=\frac{x^2-x}{5}\ ,\ x=0,1,2,3$

- ١٧ – في كل من الدوال التالية ، عين الثابت **k** بحيث تكون **f(x)** دالة كثافة احتمال :-

ا) $f(x)=\frac{k}{x^2}\ ,\ x=1,2$ ب) $f(x)=\frac{k}{x},\ x=1,2,3$

جـ) $f(x)=kx\ ,\ x=0,1,2$ د) $f(x)=k(\frac{1}{2})^x\ ,\ x=1,2,3$

ز) $f(x)=k[(\frac{1}{2})^x-\frac{1}{2}]\ ,\ x=0,1,2$

ر) $f(x)=k(8-x)\ ,\ x=0,1,2,3,4,5$

- ١٨ – إذا كان **X** متغيرا عشوائيا يمثل الزمن بالثواني الذي يستغرقه حاسب في تنفيذ برنامج ما ٠ إذا كان التوزيع الاحتمالي للمتغير **X** هو :-

$$f(x)=\frac{x}{21}\ ,x=1,2,3,4,5,6.$$

(ا) اثبت أن $f(x)$ دالة كثافة احتمال (ب) ما هو احتمال أن الزمن الذي يستغرقه الحاسب : **بالضبط 4 ثواني في التنفيذ – على الأقل 3 ثواني وليس أكثر من 5 ثواني – أكثر من 5 ثواني٠**

- ١٩ – إذا كان التوزيع الاحتمالي لعدد العملاء الذين يترددون على شركة ما في اليوم هو :-

x	P(X=x)	x	P(X=x)
5	0.02	30	0.10
10	0.05	35	0.10
15	0.15	40	0.09
20	0.20	45	0.04
25	0.25		

أوجد القيمة المتوقعة لعدد العملاء في يوم محدد٠

- ٢٠ – أى من الدوال التالية تمثل توزيع احتمالي :-

ا) $f(x)=x,\ x=\frac{1}{6},\frac{1}{3},\frac{1}{2}$ ب) $f(x)=\frac{2}{x},\ x=3,4,5$

جـ) $f(x)=\frac{x^2}{4},\ x=1,2,3$ د) $f(x)=\frac{x-3}{9},\ x=3,4,5,6,7$

$f(x) = \frac{x^2 - 1}{50},\ x = 2,3,4,5$ (ر) $f(x) = \frac{x}{3},\ x = \frac{1}{2}, 1, \frac{3}{2}$ (ز)

$f(x) = \frac{1}{x},\ x = 2,3,4,5$ (هـ)

– ٢١ – إذا كان التوزيع الاحتمالي لمتغير X هو :–

$$f(x) = \binom{3}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{3-x}, x = 0,1,2,3.$$

أوجد : (ا) توقع وتباين المتغير العشوائي X (ب) $E[\{(x - E(x)\}^2]$

(جـ) $E[(2x+1)^2]$ (د) $E(2x^2 + 6)$

– ٢٢ – الجدول الآتي يعطى الدالة الاحتمالية للمتغيرين Y,X احسب معامل الارتبـــاط وأوجـــد

$Cov(X\ ,\ 3X-7)\ ,\ Cov(2X\ ,Y)\ ,\ \sigma^2_{3X-2Y}\ ,\ E(7X-2Y)$

y \ x	0	1	h(y)
0	$\frac{6}{15}$	$\frac{4}{15}$	$\frac{10}{15}$
1	$\frac{4}{15}$	$\frac{1}{15}$	$\frac{5}{15}$
g(x)	$\frac{10}{15}$	$\frac{5}{15}$	

– ٢٣ – الجدول التالي يمثل التوزيع الاحتمالي لعدد الوحدات المعيبة في الإنتاج اليومي لأحد المصانع .

x	0	1	2	3	4	5
P(X=x)	0.884	0.1	0.01	0.003	0.002	0.001

المطلوب : (ا) تمثيل التوزيع بيانيا (ب) التباين والانحراف المعياري لعدد الوحدات المعيبة .

– ٢٤ – وعاء يحتوى على 40 كرة مرقمة من 1 إلى 40 فإذا تقرر اختيار كرة من الوعاء أذكر المتغـير العشوائي X الذي يمثل رقم الكرة المختارة من الصندوق وأوجد توقعة وتباينه .

الفصل الخامس

عرض ووصف البيانات

Presentation and Description of Data

(٥-١) المجتمعات والعينات Populations and Samples

تسجل نتيجة كل تجربة إحصائية ،كما ذكرنا في الفصل الثالث ، إما بقيمة رقمية أو تمثيل وصفي. فعلى سبيل المثال عند إلقاء زهرة نرد مره واحدة وإذا كان الإهتمام بعدد النقاط التي تظهر على السطح العلوي للنرد فإننا نسجل قيمة رقمية. بينما عند سؤال مجموعة من العاملين في هيئة ما عن الحالة الإجتماعية لكل منهم، فإن التمثيل الوصفي يكون أكثر فائدة. فالحالة الاجتماعية لأي شخص إما أعزب أو متزوج أو مطلق أو أرمل. عادة يهتم الإحصائى بالقيم الرقمية لذلك فإن التمثيل الوصفي يمكن تحويله إلى قيم عددية. فعلى سبيل المثال عند تصنيف مجموعة من الأشخاص حسب الحالة الإجتماعية فإنه يمكن تخصيص الرقم 1 للأعزب والرقم 2 للمتزوج والرقم 3 للمطلق والرقم 4 للأرمل. القيمة التي تسجل من نتيجة تجربة إحصائية تسمى بيان أو مشاهدة (مقياس) كما ذكرنا في الفصل الثالث. عندما يقوم باحث بتصنيف العاملين في شركة ما حسب الحالة الإجتماعية، في هذه الحالة يكون لديه عدد محدود من المشاهدات. بينما عند إلقاء زهرة نرد عدد لانهائي من المرات وتسجيل عدد النقط التي تظهر في كل مرة فإننا نحصل على فئة لانهائية من القيم. كل المشاهدات تحت الدراسة ، سواء كانت محدودة أو غير محدودة ، تسمى مجتمع population . في السنوات الماضية كانت كلمة مجتمع تشير إلى مشاهدات من دراسات إحصائية تشمل أشخاص. أما الآن فإن الإحصائي يستخدم هذه الكلمة لتشير إلى مشاهدات عن أي شيء موضع إهتمامه سواء مجموعة من الأشخاص، حيوانات، نباتات... الخ .

تعريف : يتكون المجتمع من كل الأشياء التي نهتم بها.

عدد المشاهدات في المجتمع تسمى حجم المجتمع وعادة يرمز لحجم المجتمع بالرمز N ، وفي هذه الحالة نقول أن المجتمع محدود. فعلى سبيل المثال عند تصنيف 500 شخصا في شركة ما حسب الحالة الإجتماعية، فإننا نقول أن المجتمع محدود وحجمه N=500. الأطوال والأوزان والدخل السنوي لمجموعة من الأشخاص أمثلة لمجتمعات محدودة. في كل حالة العدد الكلى للمشاهدات رقم محدود. في بعض الأحيان يكون حجم المجتمع غير محدود، مثل مجتمع كرات الدم البيضاء التي تسرى في دم إنسان. أيضا المشاهدات التي نحصل عليها من قياس الضغط الجوى كل يوم من الماضي إلى المستقبل تمثل مجتمع غير محدود.

كل مشاهدة في المجتمع تمثل قيمة من قيم المتغير العشوائي X ،على سبيل المثال عند إلقاء زهرة نرد عدد لانهائي من المرات وإذا كان X يمثل عدد النقط التي تظهر على النرد كل مرة،

أي أن x=1,2,3,4,5,6 ، فإن كل مشاهدة في المجتمع تمثل قيمة من قيم المتغير العشوائي X.

لمزيد من التوضيح بفرض أن المجتمع التالي يمثل عدد مرات الغياب في السنة المسجلة لكل من عشرين طالبا في كلية ما: 0,0,7,8,9,9,9,8,8,6,5,4,3,2,1,1,6,6,7,8 • وعلى ذلك X متغير عشوائي يأخذ القيم 0,1,2,3,4,5,6,7,8,9 • قيم X مع الإحتمالات المقابلة لها تعرف التوزيع الإحتمالى للمتغير العشوائي X • وبما أن الإحتمالات تمثل تكرارات نسبية، فإن المجتمع أصبح معرف تماما من توزيعه الاحتمالي• إن دراسة خواص المتغير العشوائي تكافئ دراسة خواص المجتمع. على سبيل المثال القيمة المتوقعة للمتغير X هي :-

$$E(X) = (0)(\frac{2}{20}) + (1)(\frac{2}{20}) + (2)(\frac{1}{20}) + (3)(\frac{1}{20}) + (4)(\frac{1}{20}) + (5)(\frac{1}{20}) +$$

$$(6)(\frac{3}{20}) + (7)(\frac{2}{20}) + (8)(\frac{4}{20}) + (9)(\frac{3}{20}) = \frac{107}{20} = 5.35.$$

والتي تساوى الوسط الحسابي للمجتمع μ وذلك من المعادلة التالية :-

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{107}{20} = 5.35.$$

تعريف : القيم العددية التي توصف المجتمع تسمى معالم.

يهتم الباحث بالوصول إلى إستنتاجات تخص معالم المجتمع، ولكن عادة يكون من المستحيل أو غير عملي ملاحظة كل قيم الفئة الممثلة للمجتمع. وعلى ذلك لابد من الإعتماد على فئة جزئية من قيم المجتمع لتساعدنا في الوصول إلى إستدلالات عن المعالم، وهذا يأخذنا إلى نظرية المعاينة theory of sampling .

تعريف : العينة sample هي فئة جزئية من المجتمع.

حتى يكون الإستدلال صحيح لابد من فهم العلاقة بين المجتمع والعينة. من المؤكد أن العينة سوف تمثل المجتمع لذلك لابد أن تكون غير متحيزة unbiased أى عينة عشوائية random sample .

تعريف : العينة العشوائية من الحجم n هي عينة تختار بحيث أن كل فئة جزئية حجمها n من مشاهدات المجتمع لها نفس الإحتمال في الاختيار.

قد نرغب في الوصول إلى إستنتاجات تخص نسبة الأشخاص المدخنين في بلد ما. في بعض الأحيان يكون من الصعوبة سؤال كل شخص في هذا البلد وحساب المعلمة التي تمثل نسبة

المدخنين الحقيقية. بدلا من ذلك نختار عينة عشوائية كبيرة ونحسب النسبة من العينة. هذه القيمة تستخدم في عمل بعض الإستدلال الذي يخص النسبة الحقيقية. القيمة المحسوبة من العينة تسمى الإحصاء statistic • وبما أن عينات عشوائية كثيرة يمكن إختيارها من نفس المجتمع فإننا نتوقع أن يختلف الإحصاء من عينة إلى أخرى، وعلى ذلك يعتبر الإحصاء متغير عشوائي•

تعريف : الإحصاء متغير عشوائي يعتمد فقط على قيم العينة المختارة•

عادة، يمثل قيمة أي إحصاء بحرف من الحروف اللاتينية الصغيرة• على سبيل المثال إذا أخذنا عينة عشوائية من الحجم n=3 من المجتمع الذي يمثل عدد مرات الغياب لـ 20 طالب و بفرض أن قيم العينة 6,7,8 . يعتبر الوسط الحسابي إحصاء وعادة يمثل بالرمز $\overline{X}$ • قيمة المتغير العشوائي $\overline{X}$ لهذا المثال هي $\overline{x}$، وعلى ذلك :-

$$\overline{x} = \frac{6+7+8}{3} = 7.$$

في التطبيق، قيمة الإحصاء تستخدم في تقدير قيمة معلمة المجتمع. لمعرفة مدى جودة قيمة الإحصاء (التقدير) لابد من معرفة التوزيع الاحتمالي للإحصاء والذي يسمى التوزيع العيني sampling distribution • التوزيعات لبعض الإحصاءات المفيدة سوف نتناولها في الفصل السادس.

(٥-٢) التوزيع التكراري Frequency Distribution

غالبا ما يكون التوزيع الاحتمالي لمتغير عشوائي غير معروف. تعتبر البيانات الإحصائية التي يجمعها الباحث بكميات كبيرة مفيدة جدا في دراسة سلوك المتغير العشوائي إذا تم عرضها بشكل مناسب. المعلومات الكثيرة يمكن الحصول عليها بتجميع البيانات في فئات classes وحساب عدد المشاهدات في كل فئة. مثل هذا التنظيم يسمى التوزيع التكراري• عندما نقوم بتجميع البيانات في فئات فإننا نحصل على أحسن صورة للمجتمع موضع الدراسة ولكننا في المقابل نفقد الكثير من التفصيلات عن المشاهدات في العينة. عدد المشاهدات في فئة خاصة يسمى تكرار الفئة class frequency ويرمز له بالرمز f • الجدول (٥-١) يمثل التوزيع التكراري لأطوال 22 نبات من نوع ما (المشاهدات معطاة لأقرب عدد صحيح). في هذا المثال لدينا 6 فئات وهم :60-64 , 55-59 , 50-54 , 45 -49 , 40-44 , 35-39 . يشار الى أصغر وأكبر القيم التي تقع في فئة معطاة بحدود هذه الفئة class limits • فعلى سبيل المثال للفئة 55 59 – أصغر رقم هو 55 ويمثل الحد الأدنى للفئة lower class limit وأكبر رقم هو 59 ويمثل الحد الأعلى للفئة upper class limit • وحيث أن البيانات الأصلية مسجلة لأقرب رقم

صحيح، فإن 4 مشاهدات تقع في الفئة 59-55 يمثلون كل المشاهدات في العينة التي قيمهم أكبر من أو يساوى 54.5 وأصغر من 59.5 . الأرقام 54.5 و 59.5 تسمى الحدود الفعلية **class boundaries** للفئة 59–55 . الرقم 54.5 يسمى الحد الأدنى الفعلي **lower class boundary** والرقم 59.5 يسمى الحد الأعلى الفعلي **upper class boundary**، أيضا الرقم 59.5 يسمى الحد الأدنى الفعلي للفئة التالية أي الفئة 64-60 . ويلاحظ أنه بالرغم من أن الفئات لها حدود فعلية مشتركة إلا أنه من غير الممكن أن تقع مشاهدة واحدة على أحد هذه الحدود وذلك لأن الحدود الفعلية للفئات تحتوى على خانات عشرية أكبر من تلك الموجودة في البيانات نفسها.

جدول (٥-١)

حدود الفئة	35-39	40-44	45-49	50-54	55-59	60-64
التكرار	1	2	3	4	4	8

يعرف الفرق بين الحد الأعلى الفعلي والحد الأدنى الفعلي للفئة بطول الفئة **class width** ويساوى أيضا الفرق بين الحد الأعلى والحد الأدنى للفئة زائدا وحدة دقة، أي وحدة من الوحدات التي قربت إليها الأعداد في البيانات (في هذا المثال وحدة الدقة هي الواحد الصحيح لأننا قربنا البيانات لأقرب رقم صحيح). من الناحية العملية يفضل الحصول على فئات ذات أطوال متساوية ما أمكن. سوف نرمز لطول الفئة بالرمز Δ . أطوال الفئات في جدول (٥-١) متساوية وتساوى $\Delta = 5$.

منتصف الفئة **midpoint** تسمى مركز الفئة **class midpoint** أو **class mark** ونحصل عليها بجمع الحد الأدنى الفعلي والحد الأعلى الفعلي للفئة وقسمة المجموع على 2 وذلك تحت فرض أن جميع المشاهدات داخل الفئة تأخذ قيما تتطابق مع مركز الفئة. مثال ذلك إفتراض أن 8 تكرارات في الفئة 64-60 تأخذ القيمة 62 والتي تمثل مركز هذه الفئة. أيضا يمكن الحصول على مركز الفئة بجمع الحد الأدنى والحد الأعلى للفئة وقسمة المجموع على 2 . من جدول (٥-١) مراكز الفئات هم : 62 , 57 , 52 , 47 , 42 , 37 ، يمثل جدول (٥-١) توزيع تكراري من النوع الذي نشاهده في التقارير المنشورة في الصحف. للأغراض الإحصائية يكون من الأفضل الحصول على توزيعات ذات تفصيلات أكثر، كما هو موضح في جدول (٥-٢) لنفس البيانات المعطاة في جدول (٥-١) .

جدول (٥-٢)

حدود الفئة	الحدود الفعلية الفئة	مركز الفئة	التكرار
35-39	34.5-39.5	37	1
40-44	39.5-44.5	42	2
45-49	44.5-49.5	47	3
50-54	49.5-54.5	52	4
55-59	54.5-59.5	57	4
60-64	59.5-64.5	62	8

مثال (٥-١) في عينة عشوائية حجمها n=35 من حيوانات التجارب تم تسجيل كمية مركب ما في الدم لكل حيوان كما في جدول (٥-٣). المطلوب عرض هذه البيانات في توزيع تكراري.

جدول (٥-٣)

1.1	1.2	1.4	1.1	1.2	1.5	1.6
1.2	1.6	1.5	1.8	1.9	1.8	1.7
1.4	1.5	2.1	2.2	2.2	2.3	2.4
2.5	2.6	2.7	2.8	2.9	2.9	2.9
2.9	2.8	2.8	2.7	2.7	2.6	2.6

الحل . في البداية نقرر عدد الفئات والتي سوف تتوزع فيها البيانات. عادة يفضل أن تكون عدد الفئات من 5 إلى 20 . إذا زاد عدد الفئات عن عشرين خسر الباحث البساطة التي يكسبها عادة عند وضع البيانات في التوزيع التكراري، وإذا قل عدد الفئات عن 5 فإن ذلك يؤدى إلى ضياع الكثير من التفصيلات الموجودة في البيانات. بفرض أننا قررنا أن يكون عدد الفئات 5 ، لحساب طول الفئة فإننا أولا نحسب المدى وهو عبارة عن الفرق بين أكبر وأصغر مشاهدة في العينة وعلى ذلك يكون المدى 2.9-1.1=1.8 . ثانيا نقسم المدى على عدد الفئات المقترحة أي $\frac{1.8}{5} = 0.36$ ويقرب الكسر الناتج من خارج القسمة إلى أقرب رقم عشري (لأن البيانات الخام أصلا مقاس لأقرب رقم عشري). أي أن طول الفئة سوف يكون $\Delta = 0.4$. نحدد بداية الفئة الأولى (الحد الأدنى للفئة الأولى) والذي غالبا ما يكون أصغر رقم في البيانات وهو 1.1 ، وكذلك نحدد الحد الأدنى للفئة الثانية بإضافة طول الفئة إلى الحد الأدنى للفئة الأولى، وهكذا لتعيين الحدود الدنيا لباقي الفئات. أما بالنسبة لتحديد الحد الأعلى للفئة الأولى فإنه يمكن تعيينه بإضافة طول الفئة إلى الحد الأدنى للفئة الأولى ثم نطرح من حاصل الجمع مقدار وحدة دقة من الوحدات التي قربت إليها الأعداد في المشاهدات، أي 0.1 . وكذلك نحدد الحد الأعلى للفئة الثانية بإضافة طول الفئة إلى الحد الأعلى للفئة الأولى، وهكذا لتعيين الحدود العليا لباقي الفئات،

وذلك تحت شرط أن الفئات متساوية الأطوال. أخيرا نقوم بعد عدد المشاهدات التي تقع في كل فئة ويوضع الرقم في عمود التكرار، ولابد أن يكون مجموع التكرارات مساويا لعدد المشاهدات في جدول (٥-٣) ۰ يمثل جدول (٥-٤) التوزيع التكراري للبيانات المعطاة في جدول (٥-٣). التكرار النسبي لكل فئة يمكن الحصول عليه بقسمة تكرار الفئة على مجموع التكرارات. الجدول الذي يحتوى على التكرارات النسبية يسمى التوزيع النسبي the relative frequency ۰التوزيع النسبي للبيانات المعطاة في جدول (٥-٤) موضح في جدول (٥-٥)۰ وبضرب كل تكرار نسبي في 100 نحصل على التكرار المئوي percentage distribution۰

جدول (٥-٤)

حدود الفئة	الحدود الفعلية الفئة	مركز الفئة	التكرار
1.1-1.4	1.05-1.45	1.25	7
1.5-1.8	1.45-1.85	1.65	8
1.9-2.2	1.85-2.25	2.05	4
2.3-2.6	2.25-2.65	2.45	6
2.7-3.0	2.65-3.05	2.85	10
المجموع			35

جدول (٥-٥)

حدود الفئة	1.1-1.4	1.5-1.8	1.9-2.2	2.3-2.6	2.7-3.0	المجموع
التكرار النسبي	0.2000	0.2286	0.1143	0.1714	0.2857	1

في بعض الأحيان يكون الإهتمام ليس فقط بعدد المشاهدات في فئة معطاة ولكن في عدد المشاهدات الذي يقع فوق أو تحت قيمة معينة. على سبيل المثال في جدول (٥-٤) عدد الحيوانات التي كمية المركب في دمها 2.25 أو أقل هو 19=7+8+4، والذي يمثل التكرار المتجمع cumulative frequency للفئة الرابعة. جدول (٥-٦) يوضح التكرارات المتجمعة والتي تم حسابها من جدول (٥-٤) ، ويسمى التوزيع التكراري المتجمع cumulative frequency distribution .

أما إذا استخدمنا التكرارات النسبية بدلا من التكرارات فإننا نحصل على التكرار المتجمع النسبي relative cumulative frequency. وإذا استخدمنا التكرارات المئوية بدلا من التكرارات فإننا نحصل على التوزيع التكراري المتجمع المئوي۰ كما يمكن الحصول

التوزيع التكراري المتجمع المئوي من جدول التكرار المتجمع وذلك بقسمة التكرار المتجمع لكل فئة على مجموع التكرارات وضرب الناتج في 100% • يوضح جدول (٥-٧) التوزيع التكراري المتجمع المئوي للبيانات في جدول (٥-٦)•

جدول (٥-٦)

الحدود العليا للفئات	التكرار المتجمع
أقل من 1.05	0
أقل من 1.45	7
أقل من 1.85	15
أقل من 2.25	19
أقل من 2.65	25
أقل من 3.05	35

جدول (٥-٧)

الحدود العليا للفئات	التكرار المتجمع %
أقل من 1.05	00.00
أقل من 1.45	20.00
أقل من 1.85	42.86
أقل من 2.25	54.29
أقل من 2.65	71.43
أقل من 3.05	100.00

مثال (٥-٢) البيانات في جدول (٥-٨) تمثل الدخل اليومي الصافي-لأقرب جنيه- في محل تجارى في مدة 40 يوم. كون جدول تكراري لهذه البيانات على أن يحتوى الجدول على مركز الفئة – الحدود الفعلية للفئة – التكرار المئوي – التكرارالمتجمع .

جدول (٥-٨)

118	124	128	134	135	138	140	142
125	130	136	138	141	143	145	144
144	146	147	150	152	154	155	146
146	147	155	168	157	163	160	163
146	168	170	175	181	181	175	168

الحل . أولا نحسب المدى = أكبر قيمة–أصغر قيمة أي 181-118=63 وحيث أن عدد الفئات المقترحة 13 فإن طول الفئة هو $\frac{63}{13}=4.846$ أي تقريبا 5 لأن البيانات مقاسة لأقرب عـــدد صحيح. الحد الأدنى للفئة الأولى هو 118 والذي يمثل أصغر قيمة في البيانات.جـــدول (٥–٩) يحتوى على المطلوب.

جدول (٥–٩)

حدود الفئة	الحدود الفعلية	مركز الفئة	التكرار	التكرار المئوي	التكرار المتجمع
118-122	117.5-122.5	120	1	2.5	1
123-127	122.5-127.5	125	2	5.0	3
128-132	127.5-132.5	130	2	5.0	5
133-137	132.5-137.5	135	3	7.5	8
138-142	137.5-142.5	140	5	12.5	13
143-147	142.5-147.5	145	10	25.0	23
148-152	147.5-152.5	150	2	5.0	25
153-157	152.5-157.5	155	4	10.0	29
158-162	157.5-162.5	160	1	2.5	30
163-167	162.5-167.5	165	2	5.0	32
168-172	167.5-172.5	170	4	10.0	36
173-177	172.5-177.5	175	2	5.0	38
178-182	177.5-182.5	180	2	5.0	40
المجموع			40	100	

مثال (٥–٣) أخذت عينة من مزرعة دواجن وكانت أوزان الدجاج مقاسه لأقرب مائة جـــرام كما في جدول (٥–١٠). كون جدول تكراري لهذه البيانات على أن يحتوى الجدول علـــى مركز الفئة – الحدود الفعلية للفئة – التكرار المئوي – التكرارالمتجمع .

الحل . أولا نحسب المدى = أكبر قيمة–أصغر قيمة أي 1400-600=800 وحيـــث أن عـــدد الفئات المقترحة 5 فإن طول الفئة هو $\frac{800}{5}=160$ أي تقريبا 200 لأن البيانات مقاس لأقرب 100 جرام. الحد الأدنى للفئة الأولى هو 600 والذي يمثل أصغر قيمة في البيانات. جـــــدول (٥–١١) يحتوى على المطلوب.

جدول (٥–١٠)

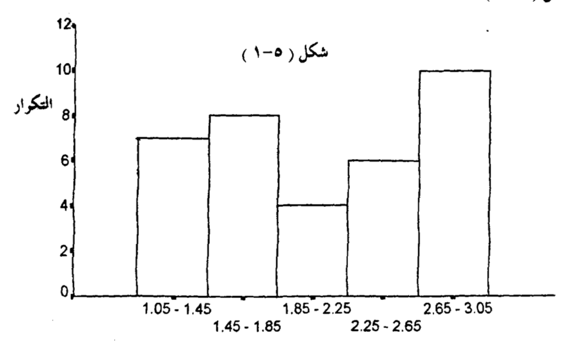
600	600	900	1000	800	800	900
700	800	800	900	700	700	700
600	900	700	1100	1200	1200	1300
1300	1300	1400	1000	1200	1200	1400
1300	1200	1300	1300	1400	1400	1000

جدول (٥-١١)

حدود الفئة	الحدود الفعلية	مركز الفئة	التكرار	التكرار المئوي	التكرار المتجمع
600-700	550-750	650	8	22.86	8
800-900	750-950	850	8	22.86	16
1000-1100	950-1150	1050	4	11.43	20
1200-1300	1150-1350	1250	11	31.42	31
1400-1500	1350-1550	1450	4	11.43	35
المجموع			35	100	

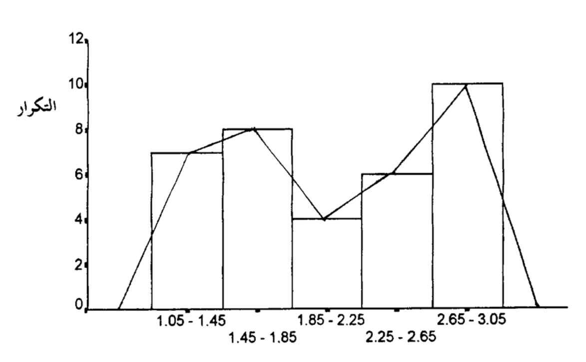
(٥-٣) التمثيل البياني **Graphic Representation**

المعلومات التي نحصل عليها من التوزيع التكراري في شكل جدولي تصبح أسهل في الفهم إذا تم عرضها بيانيا • من أكثر الأشكال البيانية الواسعة الانتشار في تمثيل البيانات الرقمية ما يعرف بالمدرج التكراري **histogram** والذي يناسب البيانات المتصلة. ونحصل عليه بتمثيل تكرار كل فئة من فئات التوزيع بمستطيل قاعدته الحدود الفعلية لتلك الفئة وارتفاعه يساوى تكرار الفئة. ويتم ذلك برسم محورين أحدهما أفقي والآخر رأسي ونرصد على المحور الأفقي الحدود الفعلية لكل فئة من فئات التوزيع التكراري ونقيم على كل فئة مستطيل ارتفاع يساوى تكرار تلك الفئة. المدرج التكراري للتوزيع التكراري المعطى في جدول (٥-٤) موضح في شكل (٥-١) •



في بعض المشاكل يكون من الأفضل وضع التكرار النسبي أو المئوي على المحور الرأسي. المدرج التكراري في هذه الحالة يسمى المدرج التكراري النسبي relative frequency histogram أو المدرج التكراري المئوي percentage frequency histogram ، ويكون لهما نفس شكل المدرج التكراري، كما أن مجموع مساحات الأعمدة للمدرج التكراري النسبي تساوى الواحد الصحيح.

الطريقة الثانية المفيدة لتمثيل البيانات الرقمية بيانيا هو استخدام المضلع التكراري frequency polygon والذي نحصل عليه بتصنيف الأضلاع العلوية للمستطيلات في المدرج التكراري ثم نوصل هذه النقاط بعضها بالبعض. ولكي نغلق الخط المنكسر الذي حصلنا عليه نحدد على المحور الأفقي مركز الفئة السابقة للفئة الأولى ومركز الفئة اللاحقة للفئة الأخيرة ونغلق المضلع. يوضح شكل (٥-٢) المضلع التكراري للتوزيع التكراري في جدول(٥-٤) .ويتضح من الشكل أن مجموع مساحات المستطيلات يساوى المساحة تحت المضلع التكراري، ويحدث هذا فقط في حالة الفئات المتساوية .

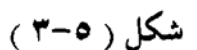


شكل (٥-٢)

وهناك طريقة أخرى لرسم المضلع التكراري وذلك برسم محورين أحدهما أفقي والآخر رأسي. يمثل المحور الأفقي مراكز الفئات ويمثل المحور الرأسي التكرارات. نعتبر مركز كل فئة إحداثيا أفقيا لنقطة ونعتبر تكرار هذه الفئة الإحداثي الرأسي لتلك النقطة. ولكي نغلق الخط

المنكسر الذي حصلنا عليه نحدد على المحور الأفقي مركز الفئة السابقة للفئة الأولى ومركز الفئة اللاحقة للفئة الأخيرة ونغلق المضلع. يوضح شكل (٥-٣) المضلع التكراري للتوزيع التكراري في جدول (٥-٤) وذلك بتحديد النقاط الآتية على الرسم :-

(1.25,7), (1.65,8), (2.05,4), (2.45,6), (2.85,10) والنقطتان الإضافيتان هما (0.85,0), (3.25,0). ثم وصل هذه النقاط بعضها بالبعض.



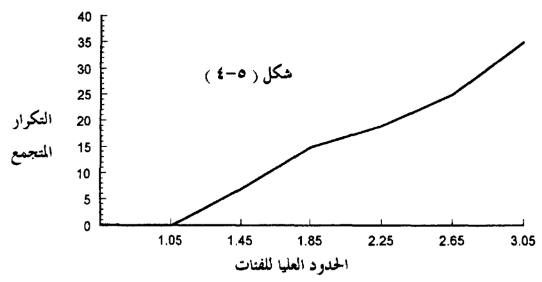
شكل (٥-٣)



عند الرغبة في مقارنة فئتين من البيانات مختلفتين في عدد مفرداتهما فإنه يمكن تمثيل المضلعين التكراريين على نفس الرسم. في هذه الحالة لابد من استخدام التكرارات النسبية أو المئوية.

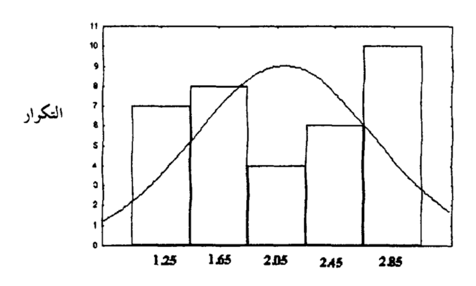
أما بالنسبة للتوزيعات التكرارية المتجمعة فهناك ما يسمى المضلع التكراري المتجمع cumulative frequency polygon ونحصل عليه برسم محورين أحدهما أفقي والآخر رأسي وكل نقطة على الرسم إحداثياها الحد الأعلى الفعلي للفئة والتكرار المتجمع وبتوصيل النقاط نحصل على المضلع التكراري المتجمع. شكل (٥-٤) يوضح المضلع التكراري المتجمع للتوزيع التكراري المتجمع في جدول (٥ - ٦).

يمكن تمثيل التوزيع التكراري المتجمع النسبي والتوزيع التكراري المتجمع المئوي بيانيا، بنفس الطريقة التي مثلنا بها التوزيع التكراري المتجمع بيانيا، وذلك باستخدام التكرارات المتجمعة النسبية والتكرارات المتجمعة المئوية.



شكل (٥-٤)

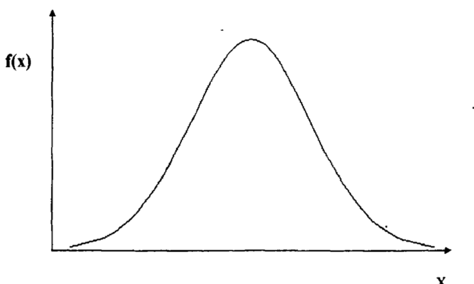
هناك طريقة أخرى لتمثيل التوزيعات التكرارية بيانيا وذلك باستخدام المنحنى التكـــراري frequency curve . ونحصل عليه برسم المضلع التكراري وتمهيد الخطوط المنكسرة التي تصل بين هذه النقط. وقد يكون التمهيد باليد أو بطرق رياضية ولا يشترط أن يمـــر المنحـــنى بجميـــع رؤوس المضلع التكراري. يبين شكل (٥-٥) المدرج التكراري والمنحنى التكراري معا للتوزيـــع التكراري في جدول (٥-٤)، عموما كلما ضاقت أطوال الفئات وزاد عدد المشاهدات فـــإن المضلع التكراري يؤول إلى المنحنى التكراري.



شكل (٥-٥)

عند الرغبة في تقدير التوزيع الاحتمالي لمتغير عشوائي متصل X نقوم بتمهيد المنحنى التكراري النسبي• شكل المنحنى يساعدنا في اقتراح شكل f(x) للمتغير العشوائي المتصل موضع الدراسة. على الرغم من أننا تمكنا من الحصول على تقدير للدالة f(x) بيانيا فلا نزال نجهل الصيغة الرياضية أو المعادلة الخاصة بالدالة f(x) وبالتالي لا نستطيع حساب تقديرات للاحتمالات. كثير من التوزيعات المتصلة يمكن تمثيلها بيانيا بمنحنى على شكل الناقوس **bell** كما في شكل (٥-٦) • الصيغة أو المعادلة الخاصة بالدالة f(x) معروفة (دالة كثافة الاحتمال) وتعتمد على معلمتين μ , σ. بمجرد الحصول على تقدير لكل منهما من البيانات يمكن كتابة المعادلة المقدرة ثم استخدام الجداول المناسبة لحساب أي مساحة تحت المنحنى•

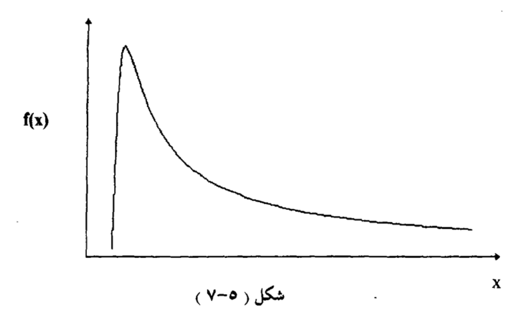
شكل (٥ -٦)



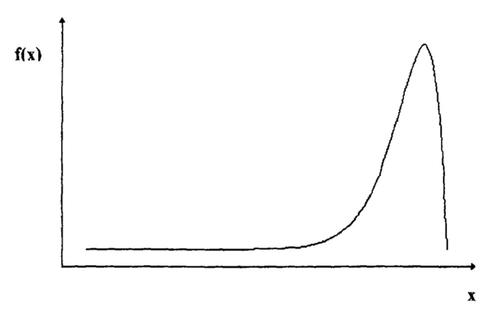
عادة تأخذ المنحنيات أشكال كثيرة ويمكن اشتقاق معادلتها بتقدير معالمها المجهولة. مشاكل التقدير صعبة جدا وتنتمي إلى فرع الإحصاء الرياضي.

يقال للتوزيع أنه متماثل **symmetrical** ، كما في شكل (٥-٦) إذا أمكننا إقامة عمود على المحور الأفقي بحيث يقسم هذا العمود التوزيع إلى قسمين ينطبقان على بعضهما تمام الانطباق. النقطة التي تمكننا من إقامة العمود تسمى نقطة التماثل. أما التوزيعات التي يكون عدم التماثل واضحا فتسمى توزيعات ملتوية **skewed** . يكون التوزيع ملتويا إلى اليمين أو موجب الالتواء **positive skewed** إذا كان معدل التناقص في المنحنى أسرع جهة اليمين منه جهة اليسار بحيث يكون الجانب الأيمن من المنحنى أطول من الجانب الأيسر كما في شكل (٥-٧

)٠بينما يكون التوزيع ملتويا إلى اليسار وسالب الالتواء negative skewed إذا كان معـــدل التناقص في المنحنى أسرع جهة اليسار منه جهة اليمين بحيث يكون الجانب الأيسر مـــن المنحـــنى أطول من الجانب الأيمن كما في شكل (٥–٨)٠



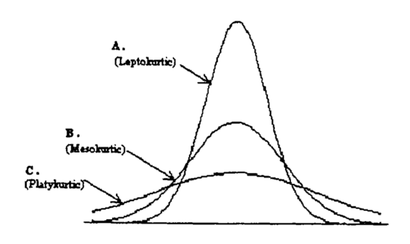
شكل (٥–٧)



شكل (٥–٨)

عند مقارنة المنحنيات وحيدة القمة قد نجدها تختلف من حيث شكل القمـــة. فقـــد تكون قمة إحداهما أكثر تدببا أو تفرطحا من بعض القمم الأخرى٠ ففي شكل (٥–٩) ثلاث منحنيات مختلفة في كمية التفلطح٠ المنحنى ، A ، ذو القمة المدببة leptokurtic يمثل توزيـــع بقيم تتركز بشدة حول نقطة الوسط midpoint ٠ المنحنى الثاني ، B ، وهو المعتدل

mesokurtic يكون متوسط التفلطح ويمثل توزيع بقيم تتركز بدرجة أقل حول نقطة الوسط عن المنحنى المدبب • وأخيرا المنحنى ، C ، المفلطح platykurtic والذي يكون منبسطا وتنخفض قمته عن قمة المنحنى المعتدل • وهذا يدل على أن قيمه تقع حول نقطة الوسط في مدى غير ضيق •



شكل (٥-٩)

(٥ - ٤) مقاييس النزعة المركزية Measures of Central Tendency

في بعض الأحيان التمثيل البياني وحده لا يمد الباحث بكل المعلومات التي يحتاج إليها من فئة المشاهدات تحت الدراسة، فالبيانات لابد أن توصف وتحلل. واحد من الطرق لوصف فئة من المشاهدات، سواء عينة أو مجتمع ، هو استخدام المتوسطات averages (مقاييس النزعة المركزية) • فالمتوسط هو القيمة التي تتركز حولها معظم المشاهدات • في هذا البند سوف نستعرض أربعة مقاييس للنزعة المركزية •

(٥-٤-١) الوسط الحسابي Arithmetic Mean

يعتبر الوسط الحسابي من أفضل مقاييس النزعة المركزية لتمثيل مركز فئة من المشاهدات.

تعريف : إذا كانت الفئة من المشاهدات $x_1, x_2, \ldots, x_N$ ، ليس من الضروري أن تكون كلها مختلفة ، تمثل مشاهدات مجتمع محدود من الحجم N ، فإن الوسط الحسابي للمجتمع يمكن حساب من الصيغة التالية :-

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}.$$

مثال (٥-٤) أوجد الوسط الحسابي لمجتمع مشاهداته هي :-

8,10,13,9,7,11,10,12,10,9,11.

الحل.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{8+10+13+9+7+11+10+12+10+9+11}{11}$$
$$= \frac{110}{11} = 10.$$

تعريف : إذا كانت الفئة من المشاهدات $x_1, x_2, \ldots, x_n$ ، ليس من الضروري أن تكون كلها مختلفة ، تمثل مشاهدات عينة من الحجم n ، فإن الوسط الحسابي للعينة يمكن حسابه من الصيغة التالية :-

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

مثال (٥-٥) أوجد الوسط الحسابي لعينة مشاهداتها هي :-

6,7,7,8.

الحل .

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{6+7+7+8}{4}$$
$$= \frac{28}{4} = 7.$$

عند وضع البيانات في توزيع تكراري فإننا نفقد الهوية لأي مشاهدة في العينة. المعلومات التي تبقى هي عدد المشاهدات التي تقع في كل فئة. لحساب الوسط الحسابي من توزيع تكراري نفترض أن كل المشاهدات داخل فئة معطاة تقع عند مركز الفئة .

تعريف : إذا كانت $x_1, x_2, \ldots, x_k$ هي مراكز الفئات لتوزيع تكراري مع تكراراتها المقابلة $f_1, f_2, \ldots, f_k$ (حيث k تمثل عدد الفئات) فإن الوسط الحسابي يحسب من الصيغة التالية :

$$\overline{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}.$$

مثال (٥-٦) أوجد الوسط الحسابي للبيانات المعطاة في جدول (٥-١٢) والتي تمثل التوزيع التكراري لأطوال عينة من الأسماك.

جدول (٥-١٢)

حدود الفئة	التكرار f_i	مركز الفئة x_i	$f_i x_i$
30-39	11	34.5	379.5
40-49	12	44.5	534.0
50-59	16	54.5	872.0
60-69	23	64.5	1483.5
70-79	17	74.5	1266.5
80-89	11	84.5	929.5
90-99	10	94.5	945.0
المجموع	100		6410

الحل • بما أن $k = 7$, $\sum_{i=1}^{k} f_i = 100$, $\sum_{i=1}^{k} f_i x_i = 6410$ وعلى ذلك الوسط الحسابي هو :

$$\overline{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i} = \frac{6410}{100} = 64.1.$$

في بعض الأحيان يكون من المناسب إضافة (أو طرح) ثابت إلى القيم الأصلية ثم حساب الوسط الحسابي. السؤال الآن كيف تكون العلاقة بين الوسط الحسابي الجديد والوسط الحسابي لفئة المشاهدات الأصلية ؟ من نظرية (٤-١) بوضع $a = 1$ ، فإن $E(X + b) = E(X) + b$ ، وعلى ذلك فإن إضافة (أو طرح) ثابت إلى كل القيم سوف يغير الوسط الحسابي بنفس المقدار • على سبيل المثال لإيجاد الوسط الحسابي للقيم $-4,-2,1,4,6$ وبفرض أننا أضفنا 4 وحصلنا على القيم الموجبة التالية : $0,2,5,8,10$ والتي وسطها الحسابي 5 . وعلى ذلك فإن الوسط الحسابي للقيم الأصلية $5 - 4 = 1$ • أيضا من نظرية (٤-١) و نتيجة (٢) فإن $E(aX) = aE(X)$ • وعلى ذلك بضرب (أو قسمة) ثابت إلى فئة من البيانات، فإن المشاهدات الأصلية سوف يكون وسطها الحسابي عبارة عن

الوسط الحسابي للمشاهدات الجديدة مقسوما على (مضروبا في) الثابت. على سبيل المثال الوسط الحسابي للقيم 5,10,15 هو 10، وعلى ذلك بعد قسمة كل المشاهدات على 5، فإن المشاهدات الجديدة هي 1,2,3 والتي وسطها 2 . وعلى ذلك الوسط الحسابي للقيم الأصلية هو $(2)(5) = 10$.

ومن مميزات الوسط الحسابي أنه مألوف وسهل الفهم كما أنه معروف لأى فئة من البيانات وقيمته وحيدة . أيضا كل قيمة في فئة المشاهدات تدخل في حسابه.

أما عيوبه فهي تأثره بالقيم الشاذة ولذلك لا ينصح باستخدامه للبيانات التي منحناها شديد الالتواء. أيضا لا يمكن تقديره من التوزيعات التكرارية التي تحتوى على فئات مفتوحة. وأخيرا لا يمكن حسابه بالرسم.

ومن الخصائص المميزة للوسط الحسابي أن مجموع انحرافات القيم (الانحراف هو بعد أي مشاهدة عن المتوسط) المشاهدة عن الوسط الحسابي يساوى صفرا، أى أن $\sum_{i=1}^{n}(x_i - \overline{x}) = 0$ ، أما في حالة التوزيعات التكرارية فإن $\sum_{i=1}^{k}(x_i - \overline{x}) \cdot f_i = 0$. أيضا مجموع مربعات انحرافات القيم حول الوسط الحسابي $\sum_{i=1}^{n}(x_i - \overline{x})^2$ أقل ما يمكن، أي أقل من مجموع مربعات انحرافات القيم عن أى قيمة أخرى غير الوسط الحسابي.

عند حساب الوسط الحسابي يفترض أن كل قيمة لها نفس الأهمية، مثل هذا الفرض قد يكون خاطئ. في الحقيقة، إذا كانت القيم ليس لها نفس الأهمية يكون من الأفضل حساب الوسط الحسابي المرجح. فإذا كانت $x_1, x_2, \ldots, x_n$ تمثل قيم المتغير X، وكانت $w_1, w_1, \ldots, w_n$ الأوزان المناظرة لها فإن الوسط الحسابي المرجح يكون :-

$$\overline{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}.$$

مثال (٥-٧) تدفع شركة أجرا قدره 25 جنيها في الساعة لعمالها غير مهرة وعددهم 20 وتدفع 30 جنيها في الساعة للعمال شبه مهرة وعددهم 10 وتدفع 40جنيها في الساعة للعمال المهرة وعددهم 5 . ما هو الوسط الحسابي المرجح للأجر في الساعة التي تدفعه الشركة ؟

الحل .

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} = \frac{(25)(20)+(30)(10)+(40)(5)}{20+10+5} = \frac{1000}{35} = 28.57.$$

يعاب على الوسط الحسابي المرجح هو عدم وجود قاعدة لتحديد الأوزان وتخضع للتقدير الشخصي. يعتبر الوسط الحسابي حالة خاصة من الوسط الحسابي المرجح إذا كانت الأوزان متساوية •

إذا كانت لدينا k من المجموعات، وكانت أحجام عيناتها $n_1, n_2, \ldots, n_k$ وأوساطها الحسابية على التوالي $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ فإن الوسط الحسابي المرجح للمجموعة التي حجمها $(n_1 + n_2 + \ldots + n_k)$ الناتجة من دمج k من المجموعات هو :-

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{\sum_{i=1}^{k} n_i}.$$

ويسمى الوسط الحسابي المرجح للأوساط الحسابية •

مثال (٥-٨) إذا كان الوسط الحسابي لأطوال 20 حيوان من نوع ما هو $\bar{x}_1 = 10$ وكان الوسط الحسابي لأطوال 30 حيوان من مجموعة أخرى من نفس النوع هو $\bar{x}_2 = 15$ أوجد الوسط الحسابي المرجح للأوساط الحسابية •

الحل .

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i} = \frac{(20)(10)+(30)(15)}{20+30} = \frac{650}{50} = 13.$$

(٥-٤-٢) الوسيط Median

يعتبر الوسيط هو المقياس الأفضل بعد الوسط الحسابي • سوف نرمز لوسيط المجتمع بالرمز $\tilde{\mu}$ ووسيط العينة بالرمز $\tilde{x}$ • الوسيط لفئة من المشاهدات مرتبة تصاعديا (أو تنازليا) هو العدد الأوسط منها إذا كان عددها فرديا وهو الوسط الحسابي للعددين الأوسطين إذا كان عددها زوجيا •

تعريف : إذا كانت $x_1, x_2, \ldots, x_n$ تمثل فئة من المشاهدات المرتبة تصاعديا (أو تنازليا)

فإن الوسيط لهذه الفئة هو العدد $x_{(\frac{n+1}{2})}$ إذا كان n فرديا وهو العدد $\frac{1}{2}[x_{(\frac{n}{2})} + x_{(\frac{n+2}{2})}]$

إذا كانت n زوجيا •

مثال (٥-٩) أوجد الوسيط للمجتمع الذي مشاهداته 10,9,8,6,7 •

الحل . بترتيب المشاهدات تصاعديا أي 6,7,8,9,10 فإن وسيط المجتمع يكون $\tilde{\mu} = 8$ •

مثال (٥-١٠) أوجد الوسيط للعينة التي مشاهداتها هي 10,9,6,1,2,7 •

الحل . بترتيب البيانات تصاعديا أي 1,2,6,7,9,10 فإن وسيط العينة يكون :

$$\tilde{x} = \frac{6+7}{2} = \frac{13}{2} = 6.50$$

ومن الخصائص المميزة للوسيط أن مجموع القيم المطلقة لانحرافات القيم عن الوسيط أقل ما يمكن. أي أن ، $\sum_{i=1}^{n}|x_i - \tilde{x}|$ أقل ما يمكن • القيمة المطلقة للرقم هي القيمة بدون إشارة، وعلى ذلك فإن 5=|5-| •

من مميزات الوسيط أنه سهل الفهم ولا يتأثر بالقيم الشاذة.كما يمكن استخدامه في التوزيعات التي تحتوى على فئات مفتوحة •

ومن عيوب الوسيط أن كل المشاهدات لا تدخل في حسابه • كما أنه مثل الوسط الحسابي ، في بعض الأحيان ، يكون قيمة صناعية **artificial** ، بمعنى عدم وجود قيمة في فئة المشاهدات ، في الحقيقة ، تمثل الوسيط.

تعريف : لأى توزيع تكراري فإن الفئة التي تحتوى على الوسيط تسمى الفئة الوسيطة **median class** •

مثال (٥-١١) أوجد الوسيط للتوزيع التكراري المعطى في جدول (٥-١٢) •

الحل • لحساب الوسيط نفترض أن البيانات تتوزع بانتظام داخل فئة الوسيط • ومن جدول (٥-١٢) نحصل على جدول (٥-١٣) . لحساب الفئة الوسيطة نحسب موقع الوسيط وهو من جدول (٥-١٣) يساوى $\frac{n}{2} = \frac{100}{2} = 50$ (حيث $\sum_{i=1}^{k} f_i = n$) سواء كان مجموع التكرارات فرديا أو زوجيا • من جدول (٥-١٣) وفي عمود التكرار المتجمع نجد أن 62 مشاهدة قيمهم أقل من 69.5 • أيضا من عمود التكرار المتجمع نجد أن 39 مشاهدة قيمهم أقل من 59.5 • أي أن الوسيط لابد أن يقع في الفئة الرابعة، أي أن فئة الوسيط هي -59.5

69.5 ، وبما أن الوسيط هو القيمة التي ترتيبها الـ 11 =39-50 داخل الفئة الوسيطية حيث 50 تمثل ترتيب الوسيط و 39 تمثل التكرار المتجمع السابق لفئة الوسيط، وتحت فــرض أن 23 مشاهدة تمثل تكرار الفئة الوسيطة موزعة بانتظام على الفئة الوسيطة التي طولها 10 وعلى ذلــك تكون المسافة من بداية الفئة الوسيطة وموقع الوسيط هي $\frac{11}{23}\cdot 10 = 4.783$ ،

جدول (٥- ١٣)

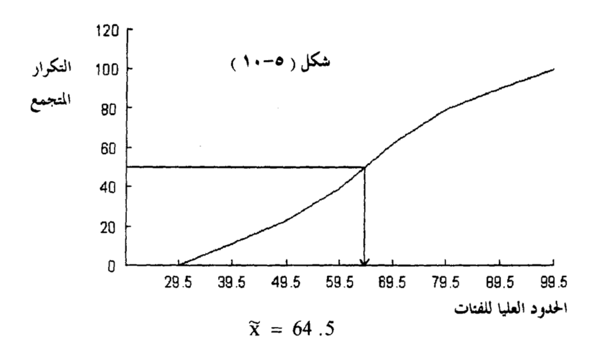
الحدود الفعلية	التكرار f_i	الحدود العليا للفئات	التكرار المتجمع
29.5-39.5	11	أقل من 39.5	11
39.5-49.5	12	أقل من 49.5	23
49.5-59.5	16	أقل من 59.5	39
59.5-69.5	23	أقل من 69.5	62
69.5-79.5	17	أقل من 79.5	79
79.5-89.5	11	أقل من 89.5	90
89.5-99.5	10	أقل من 99.5	100
المجموع	100		

أي أن الوسيط هو $\tilde{x} = 59.5 + 4.783$ أي $\tilde{x} = 64.283$ ، وعلى ذلك يمكن القــول أن 50% من الأسماك أطوالها تكون أقل من 64.283 ، ولأن الكثيرون يفضلون حساب الوســيط من صيغة فسوف نقدم لهم الصيغة التالية لحساب الوسيط :-

$$\tilde{x} = L + \left(\frac{\frac{n}{2} - F}{f_m} \right) \Delta.$$

حيث : L =الحد الأدنى الفعلي لفئة الوسيط ، F =التكرار المتجمع السابق لفئة الوســيط ، Δ =طول فئة الوسيط ، f_m = تكرار فئة الوسيط .

يمكن إيجاد الوسيط بالرسم من المضلع التكراري المتجمع. فعلى سبيل المثال للتوزيع التكراري لجدول (٥-١٣) نرسم المضلع التكراري المتجمع كما في شكل (٥-١٠) ويتم تحديد موقع الوسيط على المحور الرأسي ثم نسقط عمود من نقطة موقع الوسيط على المضلع التكراري المتجمع وعند التقائه بالمضلع نسقط عمود على المحور الأفقي فتكون هي قيمة الوسيط. من شكل (٥-١٠) فإن الوسيط يكون $\tilde{x} = 64.5$.



$\tilde{x} = 64.5$

(٥-٤-٣) المنوال **Mode**

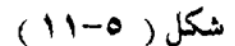
يعرف المنوال بأنه القيمة الأكثر شيوعا أو التي تتكرر أكثر من غيرها. في بعض الأحيان لا يوجد منوال لفئة من المشاهدات حيث لا تتكرر القيم أكثر من مرة ، وإذا وجد قد لا يكون وحيدا. على سبيل المثال المنوال للمشاهدات 3,5,5,5,5,5,5,5,7,7,9 هو 5 .

مثال (٥-١٢) أوجد المنوال للمشاهدات 2,4,5,6,6,6,6,6,7,7,7,7,7,8,9 .

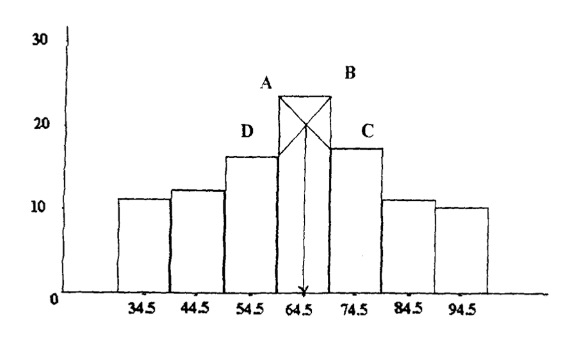
الحل. التوزيع في هذه الحالة يكون ثنائي المنوال حيث المنوال هو 6,7 .

يعتبر المنوال أقل مقاييس النزعة المركزية استخداما. للفئات الصغيرة من البيانات لا يكون له فائدة ، فقط يكون له معنى إذا كان حجم البيانات كبيرا. ومن مميزاته أنه لا يحتاج إلى عمليات حسابية، كما يمكن استخدامه للبيانات الوصفية. ويمكن حساب المنوال بالرسم من المدرج التكراري. فعلى سبيل المثال للتوزيع التكراري لجدول (٥-١٣) نرسم المدرج التكراري كما في شكل (٥-١١). نصل الرأس الأيمن العلوي للمستطيل الذي يمثل أكبر

تكرار بالرأس الأيمن للمستطيل السابق له (BD كما في الشكل) • أيضا نصل الرأس الأيسر العلوي لأطول مستطيل بالرأس الأيسر العلوي للمستطيل التالي له (AC) فيتقاطع المستقيمان في نقطة، نسقط عمود من هذه النقطة على المحور الأفقي فيكون هو المنوال • من شكل (٥-١١) المنوال يساوى 64.5 •



شكل (٥-١١)



المنوال =64.5

الحدود الفعلية للفئة

في حالة البيانات التكرارية فإن المنوال لدالة كثافة الاحتمال $f(x)$ هو قيمة x التي عندها يأخذ المنحنى أعلى قيمة • وعلى ذلك يمكن اعتبار مركز الفئة التي يقابلها أعلى تكرار هو تقدير للوسيط • وعلى ذلك للتوزيع التكراري في جدول (٥-١٢) فإن المنوال التقريبي هو 64.5 •

(٥-٤-٤) الوسط الهندسي The geometric Mean

في الحقيقة ، فإن الوسط الهندسي له استخدامات خاصة في المشاكل الاقتصادية وفي المجال السكاني •

تعريف : إذا كان لدينا الفئة من المشاهدات $x_1, x_2, \ldots, x_n$ ، فإن الوسط الهندسي يمكن حسابه من الصيغة التالية :-

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}.$$

ولتسهيل حساب الوسط الهندسي تستخدم الصيغة التالية إذا كان $n > 2$:-

$$\log G = \frac{\sum_{i=1}^{n} \log x_i}{n}$$

مثال (٥-١٣) حصل مستثمر على عائد من رأسماله المستثمر قدره 2% للسنة الأولى، 3% للسنة الثانية، 5% للسنة الثالثة أوجد الوسط الهندسي للعائد السنوى.

الحل .

$$\log G = \frac{\sum_{i=1}^{n} \log x_i}{n}$$

$$\log G = \frac{\sum_{i=1}^{n} \log x_i}{n} = \frac{1}{3}(\log 2 + \log 3 + \log 5)$$

$$\frac{1}{3}(0.3010 + 0.4771 + 0.6990)$$

$$= \frac{1.4771}{3} = 0.49237.$$

وعلى ذلك فإن $G = 3.1072$.

دائما يكون الوسط الهندسي أصغر من الوسط الحسابى . كما أن الوسط الهندسي لا يمكن حسابه إذا كانت إحدى القيم مساوية للصفر أو رقم سالب .

يمكن حساب الوسط الهندسي من جداول تكرارية من التعريف التالى .

تعريف : إذا كانت $x_1, x_2, \ldots, x_k$ تمثل مراكز الفئات لتوزيع تكراري مع تكراراها المقابلة $f_1, f_2, \ldots, f_k$ (حيث k تمثل عدد الفئات) فإن الوسط الهندسي يحسب من الصيغة التالية :

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i \log x_i}{\sum_{i=1}^{k} f_i}.$$

مثال (٥-١٤) البيانات في جدول (٥-١٤) تمثل التوزيع التكراري لأطوال عينة من خمسين نبات من نوع ما والمطلوب إيجاد الوسط الهندسى .

الحل . من جدول (٥-١٤) فإن :-

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i \log x_i}{\sum_{i=1}^{k} f_i} = \frac{55.1799}{50} = 1.103598.$$

وعلى ذلك فإن الوسط الهندسي هو $G = 12.69399$ ،

جدول (٥-١٤)

حدود الفئة	التكرار f_i	مركز الفئة x_i	$\log x_i$	$f_i \log x_i$
6-8	6	7	0.8451	5.0706
9-11	10	10	1.0000	10.0000
12-14	15	13	1.1139	16.7085
15-17	12	16	1.2041	14.4492
18-20	7	19	1.2788	8.9516
المجموع	50			55.1799

(٥-٥) الربيعات والمئينات والعشيرات

Quartiles, Percentiles, Deciles

كما ذكرنا سابقا، إذا رتبنا فئة من المشاهدات حسب قيمها تصاعديا فإن القيمة التي تكون في المنتصف والتي تقسم البيانات إلى قسمين متساويين في العدد هي الوسيط ، وبتعميم الفكرة وتقسيم البيانات إلى أربعة أجزاء متساوية (بعد ترتيب المشاهدات تصاعديا) فإن نقاط التقسيم يرمز لها بالرموز Q_3 ، Q_2 ، Q_1 حيث Q_1 يسمى الربيع الأول **first quartile** (الربيع الأدنى **lower quartile**) و Q_2 يسمى الربيع الثاني **second quartile** (نفسه الوسيط) و Q_3 يسمى الربيع الثالث **third quartile** (الربيع الأعلى **upper quartile** فالربيع الأول هو القيمة Q_1 التي يسبقها ربع البيانات ويليها ثلاثة أرباع البيانات ، والربيع الثاني (وهو أيضا الوسيط) هو القيمة Q_2 التي يسبقها نصف البيانات ويليها نصف البيانات ، وفى النهاية الربيع الثالث وهو القيمة Q_3 التي يسبقها ثلاثة أرباع البيانات ويليها ربع البيانات ، عند استخدام فئة من المشاهدات فإن الربيعات الثلاثة يتم حسابها بتعين موقعها أولا فموقع الربيع الأول هو $\frac{n+2}{4}$ والربيع الثاني موقعه هو $\frac{n+1}{2}$ والربيع الثالث موقعه هو $\frac{3n+2}{4}$ ،

للمشاهدات في شكل (٥-١٢) موقع الربيع الأول هو $\frac{12+2}{4}=3.5$ إما قيمته فهي متوسط القيمة الثالثة والرابعة أي $Q_1=\frac{5+7}{2}=6$ • أيضا موقع الربيع الثاني والثالث هما على التوالي $\frac{12+1}{2}=6.5$ ، $\frac{36+2}{4}=9.5$ وعلى ذلك فإن قيمة الربيع الثاني والثالث على التوالي هما $Q_2=\frac{11+13}{2}=12$ ، $Q_3=\frac{17+19}{2}=18$ •

شكل (٥-١٢)

1 4 5	7 10 11	13 16 17	19 20 22

عند استخدام التوزيعات التكرار يه، فإن قيمة الربيع الأول والثالث تحسب من المعادلتين:-

$$Q_3=L+\left(\frac{\frac{3n}{4}-F}{f_{Q_3}}\right)\Delta \quad ، \quad Q_1=L+\left(\frac{\frac{n}{4}-F}{f_{Q_1}}\right)\Delta.$$

حيث L = الحد الأدنى الفعلي لفئة الربيع و F =التكرار المتجمع السابق لفئة الربيع و Δ =طول فئة الربيع و f_Q =تكرار فئة الربيع•

مثال (٥-١٥) أوجد الربيع الأول والثالث للبيانات في جدول (٥-١٣)•

الحل• موقع الربيع الأول هو $\frac{n}{4}=\frac{100}{4}=25$ وفئة الربيع الأول هي **49.5-59.5** وقيمة الربيع الأول هو :-

$$Q_1=L+\left(\frac{\frac{n}{4}-F}{f_{Q_1}}\right)\Delta=49.5+\frac{(25-23)}{16}\cdot 10=49.5+1.25=50.75.$$

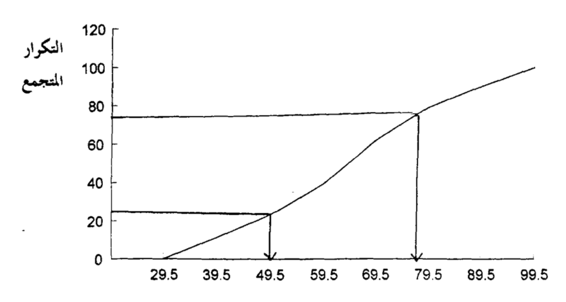
بنفس الشكل موقع الربيع الثالث $\frac{3n}{4}=\frac{(3)(100)}{4}=75$ وفئة الربيع الثالث هي - **69.5** **79.5** وقيمة الربيع الثالث هو :-

$$Q_3=L+\left(\frac{\frac{3n}{4}-F}{f_{Q_3}}\right)\Delta=69.5+\frac{(75-62)}{17}\cdot 10=69.5+7.6471=77.1471.$$

ويمكن الحصول على الربيع الأول والثالث بيانيا من المضلع التكراري المتجمع بنفس

الطريقة التي استخدمت في حساب الوسيط بيانيا مع استخدام موقع الربيع بـدلا مـن موقـع الوسيط . باستخدام البيانات المعطاة في جدول (٥-١٣) . فإن قيم الربيع الأول والثالث بيانيا هي على التوالي , 49.5 و 77.0 وذلك من شكل (٥-١٣) .

شكل (٥-١٣)



الحدود العليا للفئة $Q_1 = 49.5 \quad Q_3 = 77$

أيضا يمكن إيجاد القيم التي تقسم فئة من المشاهدات (بعد ترتيبها تصاعديا) إلى عشرة أقسـام ونرمز لنقط التقسيم بالرموز $D_1, D_2, \ldots, D_9$ حيث D_1 العشير الأول وهو يمثل القيمة التي يسبقها $\frac{1}{10}$ من البيانات ويليها $\frac{9}{10}$ من البيانات و D_2 العشير الثاني وهو يمثل القيمة الـتي يسبقها $\frac{2}{10}$ من البيانات ويليها $\frac{8}{10}$ من البيانات وهكذا للعشيرات الأخرى . بنفس الشـكل يمكن إيجاد القيم التي تقسم فئة من المشاهدات (بعد ترتيبها تصاعديا) إلى مائة قسم ونرمز لنقـط التقسيم بالرموز $P_1, P_2, \ldots, P_{99}$ حيث P_1 المئين الأول هو يمثل القيمة التي يسبقها $\frac{1}{100}$ من البيانات ويليها $\frac{99}{100}$ من البيانات و P_2 المئين الثاني وهو يمثل القيمة التي يسبقها $\frac{2}{100}$ من البيانات ويليها $\frac{98}{100}$ من البيانات وهكذا لباقي المئينات . في حالة التوزيعات التكرارية يمكـن حساب العشيرات و المئينات بنفس طريقة حساب الوسيط مع استبدال $\frac{n}{2}$ بـ $\frac{n}{10}$ للعشـير

الأول و $\frac{2n}{10}$ للعشير الثاني وهكذا لباقي العشيرات ، أيضا استبدال $\frac{n}{2}$ بـــ $\frac{n}{100}$ للمئـــين الأول و $\frac{2n}{100}$ للمئين الثاني وهكذا الباقي المئينات .

مثال (٥-١٦) أوجد العشير الأول والمئين التسعين للبيانات في جدول (٥-١٣) .

الحل . موقع العشير الأول هو $10 = \frac{100}{10} = \frac{n}{10}$ وفئة العشير الأول هي 29.5-39.5 وقيمة العشير الأول هو :-

$$D_1 = 29.5 + \frac{(10-0)}{11} \cdot 10 = 29.5 + 9.091 = 38.591.$$

حيث 29.5 = الحد الأدنى الفعلي لفئة العشير الأول و0 = التكرار المتجمـــع الســـابق لفئـــة العشير و 10= طول فئة العشير الأول و 11= تكرار فئة العشير الأول .

بنفس الشكل موقع المئين التسعين $90 = \frac{(90)(100)}{100} = \frac{90n}{100}$ وفئة المئين التسعين هي 79.5 -89.5 وقيمة المئين التسعين هو :-

$$P_{90} = 79.5 + \frac{(90-79)}{11} \cdot 10 = 79.5 + 10 = 89.5.$$

حيث 79.5= الحد الأدنى الفعلي لفئة المئين التسعين و 79= التكرار المتجمع السابق لفئة المئـــين التسعين و 10= طول فئة المئين التسعين و 11= تكرار فئة المئين التسعين .

يمكن الحصول على العشيرات والمئينات بيانيا من المضلع التكـــراري المتجمـــع بنفـــس الطريقة التي اســـــتخدمت في حســـاب الوســـيط بيانيـــا ، وممـــا يجـــدر الإشـــارة إليـــه أن $Q_2 = D_5 = P_{50}$ ، أيضا كل الربيعات والعشيرات تمثل مئينات فعلى سبيل المثال D_8 هو P_{80} و Q_3 هو P_{75} .

(٥-٦) مقاييس التشتـــت Measures of Dispersion

مقاييس النزعة المركزية التي تمت مناقشتها في البند السابق لا تكفى لإعطاء وصف كـــافي لتوزيع فئة من المشاهدات فلا توضح طبيعتها ولا كيفية توزيع مشاهداتها . كما أن الاعتماد فقط على أي مقياس للنزعة المركزية لمقارنة عدة مجموعات لا يكفى لإظهار حقيقة المقارنـــــة، فمـــن الممكن أن يكون لعدة مجموعات من البيانات نفس الوسط الحسابي والوسيط ولكنهم يختلفوا عن بعضهم تمام الإختلاف . فقد تكون مشاهدات إحدى المجموعات متقاربة بعضـــــها مـــن بعـــض (متمركزة حول متوسطها) أو مبعثرة (متشتتة) . فعلى سبيل المثال في جـــــدول (٥-١٥) ثلاثة فئات من المشاهدات A_1, A_2, A_3 كل فئة لها نفس الوسط الحسابي(60) والوسيط(60)

، ولكن يختلفوا في التشتت أو الإنتشار ، في الفئة A_1 القيم الستة لهم نفس القيمة ولا يوجد أي تشتت فهم متجانسين تماما ، في الفئة A_2 القيم تختلف من خلال ثلاث قيم وفي الفئة A_3 القيم تختلف من خلال قيمتين ولكن هناك انتشار أكثر في الفئة A_3 حيث لا يوجد قيمة في الفئة A_3 تقترب من المتوسط الحسابي أو الوسيط ، ومن هنا كان من الضروري عند وصف فئة من البيانات بمقياس رقمي أن نصفها عن طريق مقياس من مقاييس النزعة المركزية ومقياس آخر يقيس بعد البيانات عن بعضها أو بعدها عن المتوسط ، بعبارة أخرى نصف درجة تشتتها ، في الجزء التالي سوف نقدم بعض مقاييس التشتت الأكثر أهمية ،

جدول (٥-١٥)

A_1	A_2	A_3
60	35	0
60	35	0
60	60	0
60	60	120
60	85	120
60	85	120

(٥-٦-١) المدى ونصف المدى الربيعي

The range and semi interquartile range

المدى لفئة من المشاهدات، هو الفرق بين أكبر وأصغر مشاهدة ، ففي جدول(٥-١٥) المدى للفئة A_1 هو صفر، والمدى للفئة A_2 هو 50 والمدى للفئة A_3 هو 120 ، وعلى ذلك يعتبر المدى مقياس للتشتت من السهل جدا حسابه ويعطى فكرة سريعة جدا عن طبيعة البيانات ويستخدم كثيرا في مراقبة الجودة وكذلك في وصف الأحوال الجوية ، ولكن من عيوبه أنه يتأثر بالقيم الشاذة ويعطى معلومات خاطئة عن الانتشار الحقيقي لمعظم البيانات كما أنه لا يستخدم جميع البيانات في حسابه ،

عند استخدام التوزيعات التكرارية فإن المدى يحسب بعدة طرق سوف نذكر منها الطريقة التالية :

المدى= الحد الأعلى الفعلي للفئة الأخيرة–الحد الأدنى الفعلي للفئة الأولى ،

للبيانات في جدول (٥-١٣) المدى = 99.5-29.5=70 ،

هناك مقاييس أخرى للتشتت يمكن استخدامها بدلا من المدى في حالة وجود قيم شاذة ، تعتمد هذه المقاييس على إهمال جزء من البيانات عند طرفي التوزيع حتى نتخلص من القيم الشاذة وتسمى شبيهات المدى ، فمثلا بحذف أعلى 10% من المشاهدات وأصغر 10% منها

نحصل على المدى المئيني أي $P_{90} - P_{10}$ • هذا المقياس يستخدم في اختبارات الذكاء في مجـــال التربية وعلم النفس• أيضا بحذف أعلى 25% من البيانات وأصغر 25% منها نحصل على المدى الربيعي أي $Q_3 - Q_1$ • وأخيرا هناك مقياس آخر يستنتج من المدى الربيعي وهو نصف المـــدى الربيعي (الانحراف الربيعي) **semi interquartile range** ونحصل عليـــه بقســـمة المـــدى الربيعي على 2 فإذا رمزنا له بالرمز MR فإن :-

$$MR = \frac{Q_3 - Q_1}{2}.$$

وعلى ذلك فإن نصف المدى الربيعي لفئة المشاهدات في شكل (٥-١٢) هو :-

$$MR = \frac{Q_3 - Q_1}{2} = \frac{18 - 6}{2} = 6.$$

أيضا المدى الربيعي للتوزيع التكراري في جدول (٥-١٣) هو :-

$$MR = \frac{Q_3 - Q_1}{2} = \frac{77.1471 - 50.75}{2} = 13.198550.$$

يعاب على نصف المدى الربيعي أنه لا يستعمل جميع مشاهدات العينة في حسابه ولتـــلافى هذا العيب سوف نقدم في البند التالي بعض المقاييس الأخرى للتشتت التي تستخدم جميع البيانات في حسابها •

(٥-٦-٢) الانحراف المتوسط **The average Deviation**

تمثل $|x_i - \mu|$ أو $|x_i - \overline{x}|$ القيمة المطلقة لانحراف أى قيمة عن الوســـط الحســـابي للمجتمع أو العينة على التوالى •

تعريف : إذا كانت لدينا الفئة من المشاهدات $x_1, x_2, \ldots, x_n$ ، فإن الانحراف المتوسط يمكن حسابه من الصيغة التالية :-

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \overline{x}|}{n}.$$

مثال (٥-١٧) أوجد الانحراف المتوسط لفئة المشاهدات : 2,3,5,7,8

الحل . من القيم نجد أن $\overline{x} = 5$ والانحرافات : 3,2,0,2,3– والقيم المطلقـــة: 3,2,0,2,3

$-3,-2,0,2,3$ هي الانحرافات.

وعلى ذلك:

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \overline{x}|}{n} = \frac{3 + 2 + 0 + 2 + 3}{5} = 2.$$

تعريف : إذا كانت $x_1, x_2, \ldots, x_k$ تمثل مراكز الفئة لتوزيع تكراري مع تكراراهـا المقابلـة $f_1, f_2, \ldots, f_k$ فإن الانحراف المتوسط هو :-

$$M.D = \frac{\sum_{i=1}^{k} f_i |x_i - \overline{x}|}{\sum_{i=1}^{k} f_i}.$$

مثال (٥-١٨) أوجد الانحراف المتوسط للبيانات في جدول (٥-١٦) ٠

جدول (٥-١٦)

مركز الفئة x_i	التكرار f_i	$x_i - \overline{x}$	$f_i \lvert x_i - \overline{x} \rvert$
34.5	**11**	**-29.6**	**325.6**
44.5	**12**	**-19.6**	**235.2**
54.5	**16**	**-9.6**	**153.6**
64.5	**23**	**0.4**	**9.2**
74.5	**17**	**10.4**	**176.8**
84.5	**11**	**20.4**	**224.4**
94.5	**10**	**30.4**	**304**
المجموع	**100**		**1428.8**

الحل٠ الوسط الحسابي سبق حسابه من جدول (٥-١٢) وهو $\overline{x} = 64.1$ ٠وعلى ذلك فـإن الانحراف المتوسط من جدول (٥-١٦) هو :-

$$M.D = \frac{\sum_{i=1}^{k} f_i |x_i - \overline{x}|}{\sum_{i=1}^{k} f_i)} = \frac{1428.8}{100} = 14.288.$$

(٥-٦-٣) التباين <u>The Variance</u>

واحد من خصائص الوسط الحسابي ، كما ذكرنا سابقا ، هــو أن مجمـوع مربعـات انحرافات قيم المشاهدات عن الوسط الحسابي للعينة ، أي $\sum_{i=1}^{n} (x_i - \overline{x})^2$ ، أقل ما يمكن٠ وإذا كان اهتمامنا بالمجتمع فإن المقدار $\sum_{i=1}^{N} (x_i - \mu)^2$ أيضا يكون أقل ما يمكن، وعلى ذلك إذا أخذنا هذا المقدار الأخير وقسمناه على حجم المجتمع فإننا نحصل على تباين المجتمع ٠

تعريف : إذا أعطيت مجتمع محدود $x_1, x_2, \ldots, x_N$ فإن تباين المجتمع هو:-

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}.$$

الانحراف المعياري، ويرمز له بالرمز σ يمكن الحصول عليه بأخذ الجذر التربيعي للتباين.

مثال (٥-١٩) أوجد الانحراف المعياري لمشاهدات المجتمع 3,4,5,5,6,7 .

الحل . لتسهيل العمليات الحسابية يمكن استخدام جدول (٥-١٧) للحصول على الوسط الحسابي كما يأتي:-

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{30}{6} = 5.$$

أيضا التباين:-

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{6} = \frac{10}{6} = \frac{5}{3}.$$

والانحراف المعياري هو : $\sigma = \sqrt{\frac{5}{3}} = 1.29$.

جدول (٥-١٧)

x_i	$(x-\mu)$	$(x-\mu)^2$
3	-2	4
4	-1	1
5	0	0
5	0	0
6	1	1
7	2	4
30	0	10

تباين العينة ، يرمز له بالرمز s^2، ويمثل قيمة من قيم الإحصاء S^2 . عينات عشوائية مختلفة من الحجم n مختارة من نفس المجتمع، عموما ، تؤدي إلى قيم مختلفة من s^2. في كثير من المشاكل

قيمة σ^2 تكون غير معروفة وتقدر بالقيمه s^2. لكي يكون تقديرنا جيد لابد من حساب التبـــاين من صيغة بحيث في المتوسط ينتج القيمة الحقيقية σ^2. بمعنى أننا إذا أخذنا كل العينات الممكنة من الحجم n من المجتمع وحصلنا على قيمة s^2 لكل عينة، فإن متوسط كـــل قيـــم s^2 لابــد أن تساوي σ^2. الإحصاء الذي يحقق هذا الشرط يسمى إحصاء غير متحيز **unbiased**.

تعريف : إذا سحبت العينة العشوائية $x_1, x_2, \ldots, x_n$ فإن تباين العينة الغير متحيز هو :-

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}.$$

والانحراف المعياري للعينة هو $s = \sqrt{s^2}$

مثال (٥-٢٠) أوجد تباين العينة والانحراف المعياري للعينة التي مشاهداتها 12,15,17,20 .

الحل . التباين والانحراف المعياري للعينة يمكن حسابه من جدول (٥-١٨) .

جدول (٥-١٨)

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
12	-4	16
15	-1	1
17	1	1
20	4	16
64	0	34

حيث :-

الوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{64}{4} = 16.$$

وتباين العينة هو :-

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{34}{3} = 11.33.$$

والانحراف المعياري للعينة هو $\sqrt{11.33} = 3.37$.

هناك صيغة أخرى لحساب تباين العينة تفيد عند استخدام الآلة الحاسبة وهى :-

$$s^2 = \frac{1}{n-1}[\sum x_i^2 - \frac{(\sum x_i)^2}{n}].$$

مثال (٥-٢١) أوجد التباين والانحراف المعياري للعينة التي مشاهداتها هي 3,4,5,6,6,7

الحل، بما أن : $\sum_{i=1}^{n} x_i^2 = 171$, $\sum_{i=1}^{n} x_i = 31$ وعلى ذلك :-

$$s^2 = \frac{1}{5}[171 - \frac{(31)^2}{6}] = \frac{13}{6} = 2.1667.$$

الانحراف المعياري للعينة نحصل عليه بإيجاد الجذر التربيعي لتباين العينة، أي أن
$s = \sqrt{\frac{13}{6}} = 1.47196.$ ،

يمتاز الانحراف المعياري عن التباين بأنه يعبر عنه باستخدام نفس وحدات القياس كما في البيانات بينما يكون تمييز التباين " وحدات القياس مربعة " ،

عرفنا من نظرية (٤-٤) أن $\sigma^2_{X+b} = \sigma^2$ وهذا يعنى أن التباين لفئة من المشاهدات لا يتأثر إذا أضفنا ثابت أو طرحنا ثابت من كل مشاهدة، أيضا عرفنا من نظرية (٤-٥) أن $\sigma^2_{aX} = a^2\sigma^2$ وهذا يعنى أنه إذا ضربنا كل مشاهدة في ثابت (أو قسمنا على ثابت) فإن التباين الأصلي نحصل عليه من التباين الجديد بقسمته على (أو ضربه في) مربع الثابت ،

في حالة التوزيعات التكرارية فإن تباين العينة يحسب من المعادلة التالية :-

$$s^2 = \frac{1}{n-1}[\sum_{i=1}^{k} x_i^2 f_i - \frac{(\sum_{i=1}^{k} x_i f_i)^2}{n}].$$ (حيث $\sum_{i=1}^{k} f_i = n$)

مثال (٥-٢٢) أوجد التباين للمشاهدات في جدول (٥-١٩) ،

الحل، بالتعويض في صيغة التباين السابقة فإن :-

$$s^2 = \frac{1}{99}[442265 - \frac{(6410)^2}{100}] = 317.0101.$$

جدول (٥-١٩)

حدود الفئة	مركز الفئة x_i	التكرار f_i	$x_i f_i$	$x_i^2 f_i$
30-39	34.5	11	379.5	13092.75
40-49	44.5	12	534.0	23763.00
50-59	54.5	16	872.0	47524.00
60-69	64.5	23	1483.5	95685.75
70-79	74.5	17	1266.5	94354.25
80-89	84.5	11	929.5	78542.75
90-99	94.5	10	945.0	89302.50
المجموع		100	6410	442265

نظرية شيبيشي **Chebyshev's theorem**

هذه النظرية مهمة في وصف فئة من المشاهدات . إذا كانت $x_1, x_2, \ldots, x_n$ فئة من المشاهدات فإنه على الأقل $(1-\frac{1}{k^2})$ من المشاهدات سوف تقع ضمن k انحرافات معيارية من وسطها الحسابي . تعتبر النظرية غير مفيدة عندما k تساوى واحد لأن هذا يعنى أنه على الأقل $(1-\frac{1}{1^2})=0$ من المشاهدات تقع في الفترة من $\mu-\sigma$ إلى $\mu+\sigma$. عندما $k=2$ فهذا يعنى أنه على الأقل $\frac{3}{4}=(1-\frac{1}{2^2})$ من المشاهدات تقع في الفترة من $\mu-2\sigma$ إلى $\mu+2\sigma$. وأخيرا عندما $k=3$ فهذا يعنى أنه على الأقل $\frac{8}{9}=(1-\frac{1}{2^3})$ من المشاهدات تقع في الفترة من $\mu-3\sigma$ إلى $\mu+3\sigma$. بفرض أن فئة من المشاهدات تمثل عينة، وإذا كانت k=2 فالنظرية تنص على أنه على الأقل $\frac{3}{4}=(1-\frac{1}{2^2})$ من المشاهدات تقع في الفترة من $\bar{x}-2s$ إلى $\bar{x}+2s$. بنفس الشكل، النظرية تنص على أنه $\frac{8}{9}=(1-\frac{1}{2^3})$ من المشاهدات تقع في الفترة من $\bar{x}-3s$ إلى $\bar{x}+3s$.

مثال (٥-٢٣) إذا كان الوسط الحسابي والانحراف المعياري لعينة من المشاهدات هما على التوالي 60 و 10 . أستخدم نظرية تشيبيشى لوصف توزيع المشاهدات .

الحل. (ا) على الأقل $\frac{3}{4}$ من المشاهدات تقع في الفترة 60 ± 20، أى من 40 إلى 80 .

(ب) على الأقل $\frac{8}{9}$ من المشاهدات تقع في الفترة 60 ± 30، أى من 30 إلى 90 .

مثال (٥-٢٤) المشاهدات التالية تمثل مشاهدات عينة :

14,18,19,19,19,19,20,20,20,20,21,25 أثبت أنه على الأقل $\frac{3}{4}$ من المشاهدات تقع ضمن أثنين انحراف معياري من الوسط الحسابي وأنه على الأقل $\frac{8}{9}$ من المشاهدات تقع ضمن ثلاثة انحراف معياري من الوسط الحسابي .

الحل . الوسط الحسابي والانحراف المعياري هما : $\bar{x}=19.5,\quad s=2.468$ وعلى ذلك $\bar{x}\pm 2s=19.5\pm(2)(2.468)=19.5\pm 4.94$. يوضح العد الحقيقي أن الفترة من

14.56 إلى 24.44 تحتوى على 10 مشاهدة من 12 مشاهدة، تمثل تقريبا 83% من المشاهدات، أي أكثر من $\frac{3}{4}$ القيم ،أيضا $\overline{x} \pm 3s = 19.5 \pm (3)(2.468) = 19.5 \pm 7.4 = 21.96$ ، يوضح العدد الحقيقي أن الفترة من 12.1 إلى 26.9 تحتوى على كل القيم في العينة وبالتأكيد تحتوى على الأقل على $\frac{8}{9}$منهم،

الآن سوف نقدم قاعدة تصف بدقة الانتشار للتوزيع الناقوسي، يعرف التوزيع الناقوسي عادة بالتوزيع الطبيعي والذي سوف نناقشه بالتفصيل في الفصل السادس، كما تعطى هذه القاعدة نتائج جيدة للتوزيعات القريبة الشكل من التوزيع الطبيعي،

قاعدة تجريبية Empirical Rule

أعتبر توزيع فئة من المشاهدات لها التوزيع الناقوس، الفترة :-

(ا) $\overline{x} \pm s$ سوف تحتوى تقريبا على68% من المشاهدات ،

(ب) $\overline{x} \pm 2s$ سوف تحتوى تقريبا على 95% من المشاهدات ،

(جـ) $\overline{x} \pm 3s$ سوف تحتوى تقريبا على 99.7% من المشاهدات ،

مثال (٥-٢٥) المشاهدات في جدول (٥-٢٠) تمثل التوزيع التكراري لدرجات مجموعة من الطلبة في مادة الإحصاء ، والمطلوب إيجاد التوزيع التكراري ووصف البيانات منه،

الحل، للبيانات في جدول (٥-٢١)، الوسط الحسابي والانحراف المعياري على التوالي هما $\overline{x} = 66.1$ ، $s = 13.7$ وعلى ذلك يمكننا حساب الفترات:

(ا) $\overline{x} \pm s = 66.1 \pm 13.7$ ،

(ب) $\overline{x} \pm 2s = 66.1 \pm 27.4$ ،

(جـ) $\overline{x} \pm 3s = 66.1 \pm 41.1$ ،

جدول (٥-٢٠)

35	41	44	45	40	51	48
54	56	55	53	58	59	60
60	61	62	63	67	64	64
67	65	66	68	69	66	70
73	75	74	72	71	76	81
79	80	78	82	83	85	86
50	62	68	72	80	88	51
91	92					

جدول (٥-٢١)

حدود الفئة	التكرار
35-39	1
40-44	3
45-49	2
50-54	5
55-59	4
60-64	8
65-69	8
70-74	6
75-79	4
80-84	5
85-89	3
90-94	2

القاعدة التجريبية السابقة سوف تفيدنا كثيرا في وصف المشاهدات في جدول (٥-٢١) ، تبعا لهذه القاعدة فإننا نتوقع أن تقريبا 68% من المشاهدات تقع في الفترة 52.4 إلى 79.8 ، يوضح العد الحقيقي من جدول (٥-٢٠) أن 34 من 51 مشاهدة تقع في هذه الفترة وتمثل 67.0% من المشاهدات ، أيضا نتوقع أن تقريبا 95% من المشاهدات تقع في الفترة 38.7 إلى 93.5 ، يوضح العد الحقيقي أن 49 من 51 مشاهدة تقع في هذه الفترة وتمثل 96% من المشاهدات ، وأخيرا نتوقع أن تقريبا 99.7% من كل المشاهدات تقع في الفترة 25 إلى 107.2 ، العد الحقيقي يوضح أن كل المشاهدات تقع في هذه الفترة وتمثل 100% من كل المشاهدات ،

(٥-٦-٤) معامل الاختلاف Coefficient of Variation

تعتبر كل مقاييس التشتت السابقة مقاييس مطلقة لأنها تأخذ تمييز الوحدات الأصلية ولذلك لا تصلح للمقارنة بين مجموعتين وحدات القياس بينهما مختلفة ، لذلك سوف نناقش مقياس نسبي يسمى معامل الاختلاف والذي يحول الانحراف المعياري إلى مقياس نسبي باعتبار أنه نسبة مئوية من الوسط الحسابي ، ويمكن حساب معامل الاختلاف من إحدى المعادلتين التاليتين :

$$V = \left(\frac{\sigma}{\mu}\right) 100 \quad \text{أو} \quad V = \left(\frac{S}{\bar{X}}\right) 100$$

لتسهيل استخدام معامل الاختلاف نقدم المثال التالي والذي يوضح كمية الإنتاج في شركة ما خلال 80 شهرا ثم كمية الإنتاج خلال فترة ثانية مقدارها 15 شهرا وقد تم حساب الوسط الحسابي والانحراف المعياري ومعامل الاختلاف لكل مجموعة والنتائج في جدول (٥-٢٢) ، يلاحظ من النتائج أن الإنتاج خلال 15 شهرا له وسط حسابي أكبر ومعامل اختلاف أقل والذي يعتبر نتائج جيدة لمدير الإنتاج والذي يهتم بزيادة الإنتاج وانخفاض معامل الإختلاف ، وبالرغم

من أن الانحراف المعياري قد زاد من 13.2 إلى 15 إلا أنه يمكن القول بناء على معامل الاختلاف أن الفترة الثانية أقل تشتتا من الفترة الأولى٠

جدول (٥-٢٢)

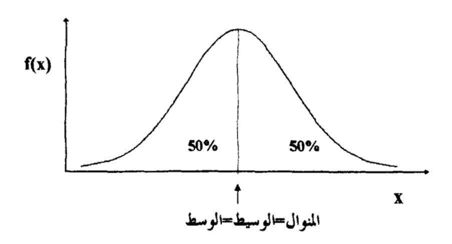
الفترة	$\bar{X}$	S	$V = (\frac{S}{\bar{X}})\,100$
80	125	13.2	10.56
15	160	15	9.375

(٥-٧) الالتواء والعلاقة بين الوسط الحسابي والوسيط والمنوال

Skewness and the Relation of the Mean , Median , and Mode

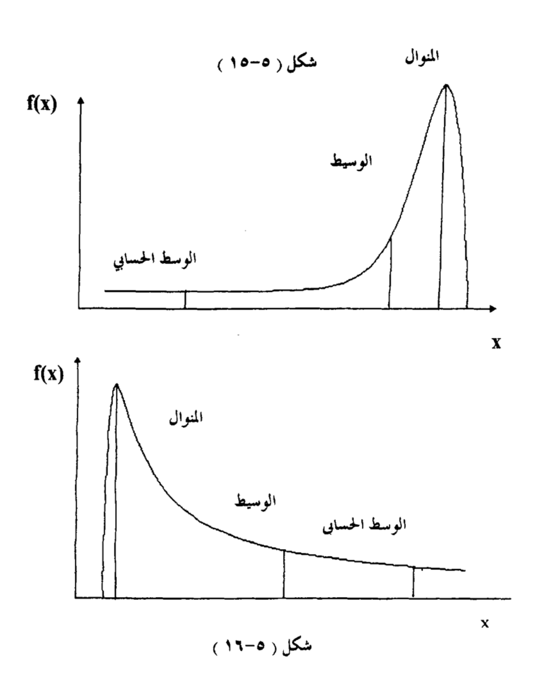
عرفنا مما سبق أن الالتواء هو بعد التوزيع التكراري عن التماثل٠ فإذا كان التوزيع متماثلا فسوف نجد أن 50% من القيم تقع على كل جانب من المنوال كما في شكل (٥-١٤)٠

شكل (٥- ١٤)



أيضا نلاحظ من شكل (٥-١٤) أن التوزيع له منوال واحد unimodal (وحيد المنوال) وأن الوسط الحسابي = الوسيط= المنوال٠ بينما في شكل (٥-١٥) نجد أن هناك علاقة بين الوسط الحسابي والوسيط والمنوال حيث الوسط الحسابي < الوسيط < المنوال وذلك لأن التوزيع ملتويا جهة اليسار٠ بينما في شكل (٥-١٦) نجد أن الوسط الحسابي > الوسيط> المنوال وذلك لأن التوزيع ملتويا جهة اليمين٠ وفى كلتا الحالتين فإن الوسيط يقع بين الوسط الحسابي والمنوال كما أن الوسط الحسابي يقع دائما في اتجاه القيم الشاذة٠

شكل (٥-١٥)



شكل (٥-١٦)

(٥-٨) بعض مقاييس الالتواء ولتفلطح

Some Measures of Skewness and Kurtosis

أولا بالنسبة لمقاييس الالتواء ، سوف نتناول مقياسين للالتواء الأول ويسمى معامل بيرسون للالتواء **Pearsonian coefficient for skewness**. تعرف معادلة معامل بيرسون للالتواء كالتالي :-

$$S_k = \frac{3(\bar{x} - \tilde{x})}{s}.$$

حيث $\overline{x}$ الوسط الحسابي و $\tilde{x}$ الوسيط و s الانحراف المعياري للعينة، ينحصر قيمة معامل بيرسون بين 3− إلى 3+ ،عندما $S_k = 0$ فهذا يعنى أن التوزيع متماثل، وإذا كانت قيمة S_k موجبة فهذا يعنى أن الوسط الحسابي أكبر من الوسيط ومن المنوال وبذلك يكون المنحنى ملتويا وله ذيل ناحية اليمين ويكون الالتواء موجبا، وأخيرا وإذا كانت قيمة S_k سالبة فهذا يعنى أن الوسط الحسابي أصغر من الوسيط ومن المنوال وبذلك يكون المنحنى ملتويا وله ذيل ناحية اليسار ويكون الالتواء سالبا ،من جدول (٥-٢٢) وإذا كان الوسيط للبيانات في الفترة التي تمثل 80 شهرا هي **115.3** فإن معامل الالتواء هو :-

$$S_k = \frac{3(\overline{x} - \tilde{x})}{s} = \frac{3(125 - 115.3)}{13.2} = +2.2045.$$

وهذا يعنى أن التوزيع به كمية من الالتواء الموجب،

يعتمد المقياس السابق للالتواء على أنه في التوزيعات الملتوية فإن الوسيط يقع تقريبا في $\frac{1}{3}$ المسافة بين الوسط الحسابي والمنوال في اتجاه الوسط الحسابي كما في شكل (٥-١٥) وشكل (٥-١٦) وهذا غير صحيح دائما، ولذلك سوف نتناول مقياس آخر للالتواء يعتمد على العزم المقدر من بيانات العينة،

تعريف : العزم r حول المتوسط لفئة المشاهدات $x_1, x_2, \ldots, x_n$ هو :-

$$m^r = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^r}{n}.$$

في حالة التوزيعات التكرارية وإذا كانت $x_1, x_2, \ldots, x_k$ تمثل مراكز الفئات لتوزيع تكراري مع تكراراتها المقابلة $f_1, f_2, \ldots, f_k$ (حيث k تمثل عدد الفئات) فإن العزم r يحسب من الصيغة التالية :-

$$m^r = \frac{\sum_{i=1}^{k} f_i (x_i - \overline{x})^r}{\sum_{i=1}^{k} f_i}.$$

المقياس الثاني للالتواء و الذي يعتمد على العزم الثالث حول الوسط الحسابي هو:

$$a_1 = \frac{m^3}{s^3}.$$

إذا كان التوزيع متماثل ، فهذا يدل على أن $a_1 = 0$ وإذا كـان $a_1 > 0$ يكـون التوزيـع موجب الإلتواء . وإذا كان $a_1 < 0$ يكون التوزيع سالب الإلتواء .

ثانيا بالنسبة لمقاييس التفلطح سوف نتناول مقياس يعتمد على العزم الرابع حول المتوسط معادلته هي :-

$$a_2 = \frac{m^4}{s^4}.$$

إذا كانت $a_2 = 3$ ، فذلك يعني أن التوزيع متوسط التفلطح . وإذا كان $a_2 > 3$ فذلك يعني أن التوزيع له قمة مدببة وإذا كان $a_2 < 3$ فهذا يدل على أن التوزيع مفلطحا .

مثال (٥-٢٦) أوجد مقيـاس الالتـواء a_1 ومقيـاس التفلطـح a_2 لفئـة المشـاهدات 2,4,6,8,13,15 .

الحل . الجدول (٥-٢٣) يسهل عملية الحساب كالتالي :-

جدول (٥-٢٣)

x_i	$(x_i - \overline{x})$	$(x_i - \overline{x})^2$	$(x_i - \overline{x})^3$	$(x_i - \overline{x})^4$
2	**-6**	**36**	**-216**	**1296**
4	**-4**	**16**	**-64**	**256**
6	**-2**	**4**	**-8**	**16**
8	**0**	**0**	**0**	**0**
13	**5**	**25**	**125**	**625**
15	**7**	**49**	**343**	**2401**
48	**0**	**130**	**180**	**4594**

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{48}{6} = 8 \quad , \quad s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} = \frac{130}{5} = 26 \quad ،$$

$$\sum_{i=1}^{n}(x_i - \overline{x})^3 = 180 \quad ، \quad s^4 = 676 \quad , \quad s^3 = (26)(5.09902) = 132.5745 \quad ،$$

وعلى ذلك العزم الثالث حول الوسط الحسابي هو :-

$$m^3 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^3}{n} = \frac{180}{6} = 30.$$

ومقياس الالتواء :-

$$a_1 = \frac{m^3}{s^3} = \frac{30}{132.5745} = 0.226288.$$

ويمكن إيجاد مقياس التفلطح a_2 وذلك بحساب القيم التالية من جدول (٥-٢٣) :-

$$\sum_{i=1}^{n}(x_i - \overline{x})^4 = 4594 \text{ ، } m^4 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^4}{n} = \frac{4594}{6} = 765.667$$

وعلى ذلك نحصل على مقياس التفلطح :-

$$a_2 = \frac{m^4}{s^4} = \frac{765.667}{676} = 1.13264.$$

وهذا يدل على أن توزيع المشاهدات مفلطح.

مثال (٥-٢٧) أوجد مقياس للتفلطح ومقياس للالتواء للمشاهدات في جدول (٥-٢٤).

جدول (٥-٢٤)

حدود الفئة	2-6	7-11	12-16	17-21	22-26	المجموع
مركز الفئة x_i	4	9	14	19	24	
التكرار f_i	2	3	5	3	7	20
$x_i f_i$	8	27	70	57	168	330
$(x_i - \overline{x})$	-12.5	-7.5	-2.5	2.5	7.5	
$(x_i - \overline{x})f_i$	-25	-22.5	-12.5	7.5	52.5	
$(x_i - \overline{x})^2 f_i$	312.5	168.75	31.25	18.75	393.75	925
$(x_i - \overline{x})^3 f_i$	-3906.3	-1265.6	-78.1	46.9	2953.1	-2250
$(x_i - \overline{x})^4 f_i$	48828.8	9492.0	195.3	117.3	22148.3	80781.7

الحل. من جدول (٥-٢٤) فإن الوسط الحسابي هو :

$$\overline{x} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} = \frac{330}{20} = 16.5$$

والتباين يساوى :

$$s^2 = \frac{\sum_{i=1}^{k}(x_i - \overline{x})^2 f_i}{(\sum_{i=1}^{k} f_i - 1)} = \frac{925}{19} = 48.68421.$$

$$s^3 = 339.6896 \ , \ s^4 = 2370.1523 \ , \ \sum_{i=1}^{k}(x_i - \overline{x})^3 f_i = -2250.$$

$$m^3 = \frac{\sum_{i=1}^{k}(x_i - \overline{x})^3 f_i}{(\sum_{i=1}^{k} f_i)} = \frac{-2250}{20} = -112.5.$$

وعلى ذلك يمكن الحصول على مقياس الالتواء كما يلي :-

$$a_1 = \frac{m^3}{s^3} = \frac{-112.5}{339.6896} = -0.331185.$$

ويمكن إيجاد مقياس التفلطح a_2 وذلك من جدول (٥-٢٤):

$$\sum_{i=1}^{k}(x_i - \overline{x})^4 f_i = 80781.7,$$

$$m^4 = \frac{\sum_{i=1}^{k}(x_i - \overline{x})^4 f_i}{\sum_{i=1}^{k} f_i} = \frac{80781.7}{20} = 4039.085.$$

وعلى ذلك يمكن الحصول على مقياس للتفلطح كما يلي :-

$$a_2 = \frac{m^4}{s^4} = \frac{4039.085}{2370.1523} = 1.704146.$$

تمارين

- ١ - إذا أعطيت التوزيع الاحتمالي التالي :

x	1	2	3
P(X=x)	0.25	0.25	0.5

(ا) أوجد قيم فئة المشاهدات من الحجم $N = 100$ التي مثلت هذا التوزيع .

(ب) أوجد الوسط الحسابي μ للمجتمع بطريقتين.

- ٢ - أوجد المعلمة μ للمجتمع 4,6,5 ثم ضع قائمة بكل العينات الممكنة من الحجم $n = 2$ مع إمكانية تكرار القيمة الواحدة في العينة ، واحسب لكل عينة القيمة $\bar{x}$ للإحصاء $\bar{X}$ و أثبت أن $E(\bar{X}) = \mu$ •

- ٣ - في تمرين - ٢ - أحسب s^2 لكل عينة وذلك تحت فرض (ا) μ معروفة (ب) μ غير معروفة وأثبت أن $E(S^2) = \sigma^2$ لكل حالة •

- ٤ - المطلوب إلقاء 10 عملات 100 مرة وتسجيل قيم x التي تمثل عدد مرات ظهور الصورة ثم إيجاد المدرج التكراري للتوزيع التكراري الذي يمثل عدد مرات ظهور الصورة •

- ٥ - أوجد للفئات التالية الحدود الفعلية للفئة ومركز الفئة وطول الفئة :

(ا) 8-12 (ب) 11.5-16.7 (جـ) 2.2-3.3 (د) 0.345-0.416 •

(ز) 77.45-86.12 •

- ٦ - فيما يلي التوزيع التكراري للدرجات التي حصل عليها 80 شخصا في دورة تدريبية في مجال الحاسب الآلي استمرت 6 شهور والمطلوب إيجاد : (ا) عدد الفئات (ب) طول الفئة (جـ) مركز الفئة (د) التكرار النسبي والتكرار المئوى •

حدود الفئة	50-59	60-69	70-79	80-89	90-99
التكرار	5	10	50	10	5

- ٧ - إذا كانت مراكز الفئات للتوزيع التكراري الذي يمثل أوزان مجموعة من النباتات (بالرطل) هي : 17,20,23,26,29,32,35 • أوجد حدود الفئة والحدود الفعلية وطول الفئة لهذا التوزيع •

-٨- في عينة من الزجاجات سعة كل منها لترا واحدا تم قياس ما تحتويه من سائل بالملليلتر وتم وضع البيانات في الجدول التالي:-

حدود الفئة	900-909	910-919	920-929	930-939	940-949
التكرار	5	8	8	24	15

أوجد : الوسط الحسابي لما تحتويه الزجاجات من سائل •

- ٩ - في عينة عشوائية من 20 طالب في كلية ما تم تسجيل عدد أيام الغياب لكل طالب خلال الفصل الدراسي الأول وكانت كالتالي : 1,0,3,4,5,4,1,1,1,1,1,0,0,0,0,3,3,2,2,1

أحسب كلا من :

الوسط الحسابي - الوسيط - المنوال - الانحراف المتوسط - الانحراف المعيارى •

– ١٠ – فيما يلي السرعة (بالأميال لكل ساعة) التي سجلها رادار على الطريـــق الزراعـــي لعينة عشوائية من 50 سيارة مرت عند نقطة المراقبة خلال ساعة :

60	66	54	54	49	74	71	65	56	47
59	70	71	66	65	70	65	64	63	55
55	54	60	63	61	65	45	53	54	61
54	54	53	48	47	70	74	63	62	61
54	66	70	64	65	64	63	68	66	70

(ا) ضع هذه البيانات في شكل توزيع تكرارى.

(ب) أوجد التوزيعات التكرارية النسبية والمئوية والمتجمعة.

(جـ) ارسم المدرج التكراري والمضلع التكراري لهذا التوزيع.

(د) ارسم المنحنى التكراري النسبي والمنحنى التكراري المتجمع النسبى.

– ١١ – إذا كان عدد أسماك السالمون التي تم صيدها بواسطة 10 صيادين في اليوم الأول مـــن الموسم هي : 5,6,7,7,7,8,9,10,3,7 أوجد : الوسيط – الانحـــراف المتوســـط – الانحـــراف المعيارى.

– ١٢ – الجدول التالي يمثل التوزيع التكراري لأعمار مجموعة من أباء طلبة كلية

حدود الفئة	35-39	40-44	45-49	50-54	55-59
التكرار	6	15	40	25	14

والمطلوب :

(ا) رسم المدرج التكراري والمضلع التكراري لهذا التوزيع.

(ب) إيجاد التكرار النسبي لهذا التوزيع ورسم مضلعة التكراري.

(جـ) إيجاد التكرار المئوي لهذا التوزيع ورسم مضلعة التكراري.

(د) إيجاد الوسط الحسابي والانحراف المعيارى.

(ز) ما نسبة الآباء الذين أعمارهم : ضمن انحراف معياري واحد من الوســـط الحســـابي أي الواقعة في الفترة $\bar{x} \pm s$.

(ر) كرر (ز) للفترة $\bar{x} \pm 2s$ وأيضا للفترة $\bar{x} \pm 3s$.

– ١٣ – اختيرت عينة عشوائية من 10 مسامير لتقدير كمية الضغط الضروري لكسر المســـمار وكانت النتائج كالتالي : 18,22,26,25,27,26,19,17,22,20 أحسب كلا من :

الوسط الحسابي – الوسيط – المنوال – الانحراف المتوسط – الانحـــراف المعيـــاري – مقيـــاس الالتواء a_1 ومقياس التفلطح a_2 – مقياس الالتواء لبيرسون.

– ١٤ – فيما يلي التوزيع التكراري لعدد الدقائق التي يتأخرها مجموعة من الطلبة في دخولهم المحاضرة بعد دخول الأستاذ :

التأخير بالدقائق	0	1	2	3	4	5	6
التكرار	180	1	2	3	5	6	6

أوجد الوسط الحسابي و الانحراف المعياري لهذا التوزيع .

– ١٥ – تمتلك شركة ما 10 قوارب للصيد، قامت الشركة بتسجيل تكاليف صيانة كل قارب (بالدولار) وكانت كما يلي : 500,505,460,470,530,506,994,880,600,460

أوجد :

(ا) الوسيط والوسط الحسابي والمنوال .   (ب) الربيع الأول والربيع الثالث .

(ج) الانحراف الربيعي والمدى .

(د) مقياس للالتواء وآخر للتفلطح .

- ١٦ – الجدول التالي يمثل التوزيع التكراري لتكاليف تجهيز عينة عشوائية من نبات للتصدير .

حدود الفئة	1.00-1.02	1.03-1.05	1.06-1.08	1.09-1.11	1.12-1.14
التكرار	5	25	57	40	41

المطلوب إيجاد (ا) الوسيط والوسط الحسابي والمنوال .   (ب) الربيع الأول والربيع الثالث .

– ١٧ – قامت شركة متخصصة في بيع البرامج الجاهزة بتسجيل حصيلة البيع الشهري بالدولار وذلك في مدة 28 شهرا وكانت المشاهدات كما يلي :

757.34	990.16	118.01	871.54	858.29	820.54	710.99
1150.5	1280.3	723.06	876.09	1230.9	657.98	1018.6
1140.6	997.05	657.90	999.98	1140.8	800.75	345.89
1234.8	1280.3	723.06	887.09	1209.0	670.01	678.98

(ا) كون توزيع تكراري مناسب   (ب) أوجد الوسط الحسابي لهذا التوزيع .

(جـ) أوجد الوسيط لهذا التوزيع حسابيا وبيانيا .

(د) ارسم المدرج التكراري وأحسب من المنوال .

–١٨ – إذا كانت درجة الحرارة المئوية في إحدى المدن خلال أيام إحدى الأسابيع هي :

22,9,13,12,18,15,9 أحسب ما يلي : (ا) الوسيط والوسط الحسابي والمنوال .

– ١٩ – فيما يلي 25 قيمة لمتغير عشوائي :–

34	50	44	54	33
34	32	34	27	22
25	12	60	50	13
40	34	35	35	34
23	34	6	7	18

(ا) كون توزيع تكراري مناسب •

(ب) ارسم المدرج التكراري والمضلع التكراري لهذا التوزيع •

(جـ) أوجد العزوم الأربعة الأولى حول الوسط الحسابي •

– ٢٠ – مجتمع له وسطه الحسابي $\mu = 11$ وانحرافه المعياري $\sigma = 2$ • باستخدام نظرية تشيبيشي أوجد (ا) نسبة القيم التي تقع بين 7 و 15 (ب) نسبة القيم التي تقع بين 5 و 17 •

– ٢١ – فيما يلي التوزيع التكراري للضرائب التي تم تحصيلها من مجموعة من الموظفين في شركة لتسخين البترول في عام 1990 •

حدود الفئة	400-500	501-601	602-702	703-803	804-904
التكرار	17	25	29	25	28

(ا) ارسم المضلع التكراري المتجمع ثم من الرسم حدد الوسيط والربيع الأول والربيع الثالث •

(ب) أحسب مقياس للنزعة المركزية •

(جـ) ما هو الرقم الذي يقسم التوزيع بحيث يكون عدد العاملين التي الضرائب المستحقة عليهم الأكبر منه يساوى عدد العاملين الأقل منه ؟

(د) ما نسبة العاملين الذين يدفعون ضرائب أقل من 702.5 •

– ٢٢– المشاهدات التالية تمثل التوزيع التكراري للأجور الشهرية (بالدولار) لمجموعة من العمال في شركة لصناعة الغاز الطبيعي •

الأجور	التكرار
900-950	9
951-1001	15
1002-1052	21
1053-1103	25
1104-1154	27
1155-1205	19
1206-1256	14
1257-1307	9
1308-1358	6

(ا) كون توزيع تكراري مناسب •

(ب) ارسم المدرج التكراري والمضلع التكراري لهذا التوزيع •

(جـ) أوجد العزوم الأربعة الأولى حول الوسط الحسابي •

-٢٣ - الآتي يمثل أعمار 30 عاملة في مصنع لتعبئة الحلوى :

48	30	20	21	58	43
20	34	29	61	32	50
25	32	31	50	49	24
41	35	34	47	21	23
42	33	32	34	36	34

المطلوب إيجاد التوزيع التكراري ورسم المدرج و المضلع و المنحنى التكراري •

-٢٤ - الجدول التالي يمثل التوزيع التكراري لأعمار مجموعة من الذكور والإناث في كلية ما •

العمر	15-16	17-18	19-20	21-22	23-24
الإناث	100	125	130	160	190
الذكور	110	131	150	146	200

(ا) ارسم المضلع التكراري لكل من الذكور والإناث •

(ب) أوجد الوسط الحسابي والانحراف المعياري لكل من الذكور والإناث •

(جـ) أوجد معامل الاختلاف لكل من الذكور والإناث وأي المجموعتين أكثر تشتتا •

- ٢٥ - إذا كان معدل التضخم لشعب ما هو 3% في السنة الأولى و4% في السنة الثانية و8% في السنة الثالثة • أوجد الوسط الهندسي لمعدلات التضخم •

- ٢٦ - حصل مستثمر على عائد على رأس ماله قدره 3% للسنة الأولى و5% للسنة الثانية 12% للسنة الثالثة أوجد الوسط الهندسي للعائد السنوى •

- ٢٧ - يشترى شخص 4 قمصان من الشركة A بسعر الواحد 22$ و4 قمصان من الشركة B بسعر الواحد 25$ و7 قمصان من الشركة C بسعر الواحد 30$ • أوجد متوسط سعر القميص •

- ٢٨ - أسرة لديها 8 أطفال، أعمارهم كالتالي : 8,10,6,14,14,12,18,20 أوجد :

(ا) مقاييس النزعة المركزية لهذه البيانات •

(ب) المدى - المدى الربيعي - الانحراف المتوسط - الانحراف المعيارى •

– ٢٩ – أشترى شخص 10 كيلو من السمك من نوع (A) بسعر الكيلو 4 جنيها و 5 كيلو من السمك من نوع (B) بسعر الكيلو 2.5 جنيها و3 كيلو من السمك من نوع (C) بسعر الكيلو 1.5 ، باستخدام الوسط الحسابي المرجح أوجد متوسط سعر السمك الذي أشترى به.

– ٣٠ – يقوم رجل أعمال بتأجير الشقق التي يمتلكها شهريا بأسعار مختلفة. فالشقة من النوع الممتاز بسعر 1500 جنيها والشقة من النوع الفاخر بسعر 1000 جنيها والشقة من النوع المتوسط بسعر 500 جنيها. فإذا كان الرجل يمتلك 50 شقة من النوع الممتاز و 30 شقة من النوع الفاخر و 15 شقة من النوع المتوسط. أوجد الوسط الحسابي للإيجار الشهري الذي يتحصل علية.

– ٣١ – لدى رجل أعمال أربعة حسابات في بنك ما. ففي الحساب A له 4,000 دولار وفي الحساب B له 3,000 دولار وفي الحساب C له 10,000 دولار. فإذا كان الحساب A يعطي له أرباح قدرها 5.5% كل سنة والحساب B يعطي له أرباح قدرها 4% كل سنة والحساب C يعطي له أرباح قدرها 6% كل سنة. أوجد الوسط الحسابي المرجح للربح السنوى.

– ٣٢ – فيما يلي توزيع عدد المشاريع المنفذة شهريا خلال عام 1995 في شركة بترول:

15,11,7,6,8,10,12,6,8,9,6,13 أحسب :

(ا) الوسط الحسابي – الوسيط – المنوال – الانحراف المعياري – المدى – المدى الربيعي.

– ب– مقياس الالتواء a_1 ومقياس للتفلطح a_2 – مقياس الالتواء لبيرسون.

– ٣٣ – يقوم العاملين في شركة صغيرة بالتوقيع في صفحات الزمن للدلالة على زمن مغادرة الشركة. المشاهدات التالية تمثل أزمنة المغادرة للعاملين وذلك في يوم عشوائي :

5:15	3:50	1:10	5:40	5:30	5:33
4:45	5:59	4:30	5:12	5:16	4:30
5:30	2:40	5:40	3:30	3:33	3:30
5:30	4:56	3:22	5:50	5:41	4:30
5:30	4:23	4:32	5:44	5:35	5:12

(ا) أوجد الوسط الحسابي لهذا التوزيع.

(ب) أوجد الوسيط لهذا التوزيع .

(جـ) أوجد الانحراف المعيارى لهذا التوزيع.

–٣٤ – فيما يلي توزيع عدد مرات احتراق الطعام خلال أسبوع لعينة من 10 سيدات: 3,2,4,1,1,0,0,0,4,1 والمطلوب إيجاد الوسط الحسابي والانحراف المعياري ووصف البيانات باستخدام نظرية تشيبيشى.

– ٣٥ –الجدول التالي يمثل التوزيع التكراري للودائع في بنك ما في نهاية عام 1993 :

حجم الودائع	عدد العملاء
0-500	7100
501-1001	8300
1002-1502	9345
1503-2003	9945
2004-2504	6257
2505-3005	2003
3006-3506	14
3507-4007	1445

أوجد : الوسط الحسابي – المنوال – الانحراف المتوسط – الانحراف المعياري

– ٣٦ – إذا كانت جامعة ما تمثل مجتمع عناصره عدد المدرسين في كل كلية ، بفرض أن الوسط الحسابي لعدد المدرسين في الكلية الواحدة هو $\mu = 16$ بانحراف معياري $\sigma = 15$ ، أستخدم نظرية تشيبيشى في وصف هذا المجتمع ،

– ٣٧ – أجرى اختبار في مادة الإحصاء لمجموعة من الطلبة عددهم 400 وقد وجد أن الوسط الحسابي هو $\bar{x} = 77$ والانحراف المعياري $\sigma = 5$ أستخدم نظرية تشيبيشى لوصف هذه الفئة من البيانات ،

– ٣٨ – وجد أن تلوث البحار يؤدى إلى نمو أنواع مختلفة من البكتريا ، فإذا كان عدد البكتريا لكل 100ملليمتر في 10 أماكن من مياه البحر هي : 49,69,60,41,69,70,51,68,67,66 أوجد: الوسط الحسابي – الوسيط – المنوال – الانحراف المتوسط – الانحراف المعيارى – مقياس للالتواء وآخر للتفلطح ،

– ٣٩ – تم سؤال عينة عشوائية من 10 عمال عن المسافة (بالأميال) التي يقطعونها في الذهاب إلى المزرعة التي يعملون بها وكانت إجابتهم كما يلي : 25,6,1,2,4,8,5,6,5,4 ، أوجد : الوسيط والربيع الأول والربيع الثالث والمدى الربيعى– مقياس للالتواء وآخر للتفلطح ،

– ٤٠ – المشاهدات التالية تمثل عدد المرضى الذين يتم الكشف عليهم يوميا في مستشفى خاص من قبل 10 أطباء : 15,8,6,9,15,18,21,39,5,7 : أوجد الوسط الحسابي والانحراف المعياري واستخدم نظرية تشيبيشى لوصف هذه الفئة من المشاهدات وأوجد مقياس للالتواء وآخر للتفلطح ،

– ٤١ – في جامعة ما يتقاضى الأستاذ في المتوسط $15,000 دولار بانحراف معيـــاري $5,000 بينما في جامعة أخرى يتقاضى الأستاذ أجر قدره $10,000 بانحراف معيـــاري $3,000 أوجـــد معامل الاختلاف لكل جامعة وأي الجامعتين أكثر تشتتا ؟

– ٤٢ – تمتلك شركة أربع مزارع ، بعض البيانات عن هذه المزارع في الجدول التالي :

المزرعة	عدد العمال	متوسط الأجر السنوي	الانحراف المعياري
1	100	$8,300	$75
2	180	11,200	1200
3	200	9,000	900

أوجد معامل الاختلاف لكل مزرعة وما هي المزرعة التي لها أكبر تشتت في الأجر السنوى ؟

– ٤٣ – إذا كان الوسط الحسابي للأجر الشهري للموظف في شركتين B , A متساوي وهـــو $10,000 بانحراف معياري $500 للشركة A وانحراف معياري $550 للشركة B ، أوجـــد معامل الاختلاف لكل شركة وأي الشركتين أكثر تشتتا في الأجر ؟

– ٤٤ – قام مزارع بوزن نوعين من ثمار البرتقال في المزرعة التي يمتلكها وقـــد حصـــل علـــى المشاهدات التالية : $\overline{x}_1 = 15$, $s_1 = 1$, $\overline{x}_2 = 14$, $s_2 = 2$ أوجد معامل الاختلاف لكل نوع وأي النوعين أكثر تشتتا ؟

الفصل السادس

بعض التوزيعات الاحتمالية

Some Probability Distributions

بالرغم من وجود أنواع لانهائية من التوزيعات الاحتمالية فإن عدد محدود منهم يستخدم في مجال واسع من التطبيقات الإحصائية . يتناول هذا الفصل بعض هذه التوزيعات .

(٦-١) التوزيع المنتظم **Uniform Distribution**

يعتبر التوزيع المنتظم من أبسط التوزيعات الاحتمالية المنفصلة، حيث أن جميع قيم المتغير العشوائي X لها الاحتمال نفسه . لهذا التوزيع بعض التطبيقات المحدودة خاصة في المعاينة الإحصائية .

تعريف : إذا كان فراغ المتغير العشوائي X هو $\Omega = \{x_1, x_2, \ldots, x_c\}$ فإن التوزيع المنتظم يأخذ الصيغة التالية :

$$f(x;c) = \frac{1}{c} \ , \ x = x_1, x_2, \ldots, x_c.$$

سوف نستخدم الصيغة $f(x;c)$ بدلا من $f(x)$ لتوضيح أن التوزيع المنتظم يعتمد على المعلمة c .

مثال (٦-١) يتكون الكتاب الخاص بدائرة المعلومات البريطانية لعام ما من 20 جزء، فإذا كان المطلوب اختيار جزءاً عشوائيا . أوجد التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل رقم الجزء المختار .

الحل . عند اختيار جزءاً عشوائيا من 20 جزء فإن كل عنصر في فراغ المتغير العشوائي $\Omega = \{x_1, x_2, \ldots, x_{20}\}$ يقع باحتمال $\frac{1}{20}$ وعلى ذلك يكون لدينا توزيع منتظم بدالة كثافة احتمال :

$$f(x;20) = \frac{1}{20} \ , \ x = 1,2,\ldots,20.$$

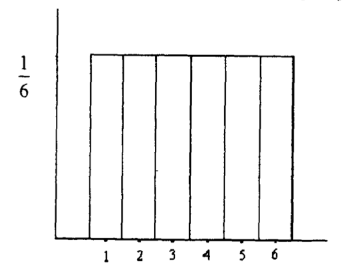
مثال (٦-٢) أوجد التوزيع الاحتمالي لكل العينات الممكن اختيارها من الحجم $n = 2$ من القيم $\{1,2,3,4\}$.

الحل . عدد العينات الممكن اختيارها هو $\binom{4}{2} = 6$ وهي على التوالي $\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}$. كل العينات السابقة لها نفس الفرصة في الظهور عند اختيار عينة عشوائية من الحجم $n = 2$ من القيم $\{1,2,3,4\}$. توزيع العينات يتبع التوزيع المنتظم بدالة كثافة احتمال :

$$f(x;6) = \frac{1}{6} \ , \ x = 1,2,\ldots,6.$$

حيث 1 تعنى العينة $\{1,2\}$ و 2 تعنى العينة $\{1,3\}$...الخ •وعلى ذلك فإن احتمال اختيار العينة $\{1,2\}$ هو $P(\{1,2\}) = P(X=1) = \frac{1}{6}$ • عموما عند اختيار عينة عشوائية من الحجم n من مجتمع محدود حجمه N فإن قيمة c في صيغة التوزيع المنتظم تعطى من $\binom{N}{n}$ •

عند التمثيل البياني للتوزيع المنتظم بالمدرج histogram نحصل على مستطيلات متساوية الارتفاع كما في شكل (٦-١) لمثال (٦-٢)•



شكل (٦-١)

(٦-٢) توزيع ذي الحدين Binomial Distribution

في كثير من الأحيان قد تشتمل تجربة ما على n من المحاولات المتكررة المستقلة بحيث يكون لكل محاولة نتيجتين اثنتين فقط، تسمى الأولى نجاح وتسمى الأخرى فشل، حيث احتمال النجاح p واحتمال الفشل $q = 1 - p$ • تسمى التجربة التي تحقق هذه الشروط بتجربة ثنائي الحدين binomial experiment • فعلى سبيل المثال عند إلقاء عملة متزنة 5 مرات حيث كل محاولة قد تكون صورة أو كتابة وذلك تحت فرض أن النجاح هو ظهور الصورة• هنا المحاولات المتكررة مستقلة كما أن احتمال النجاح ثابت من محاولة إلى أخرى ويساوى $p = \frac{1}{2}$ • ويجب ملاحظة أنه يمكن تعريف النجاح والفشل عكس ذلك تماما، أى جعل ظهور الكتابة نجاح، وفى هذه الحالة تتبدل قيمتي q , p •

عموما يمكن القول أن تجربة ذي الحدين هي التجربة التي تحقق الشروط الآتية :-

- ا - التجربة التي تتكون من n من المحاولات المتكررة.
- ب - نتيجة كل محاولة يمكن تصنيفها إلى نجاح أو فشل.
- جـ - احتمال النجاح، وهو p يبقى ثابت من محاولة إلى محاولة.
- د - المحاولات المتكررة مستقلة بعضها عن بعض.

ففي حالة إلقاء عملة 3 مرات وتحت فرض أن النجاح هو ظهور الصورة تعتبر عـــدد حـــالات النجاح متغير عشوائي يأخذ قيم صحيحة من 0 إلى 3 . النواتج الثمانية الممكنة وقيم x المقابلـــة معطاة في الجدول التالي.

النواتج	TTT	THT	TTH	HTT	THH	HTH	HHT	HHH
x	0	1	1	1	2	2	2	3

وحيث أن المحاولات مستقلة باحتمال نجاح ثابت ويساوى $\frac{1}{2}$ ، فإن :

$$P(HTH) = P(H)P(T)P(H) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{2}) = \frac{1}{8}.$$

وبنفس الشكل، تحدث النواتج الأخرى باحتمال $\frac{1}{8}$. التوزيع الاحتمالي للمتغير العشـــوائي X معطى في الجدول التالي :

x	0	1	2	3
P(X=x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

هذا و يمكن وضع صيغة للتوزيع الاحتمالي للمتغير العشوائي X على الشكل :-

$$f(x) = \frac{\binom{3}{x}}{8}, \quad x = 0,1,2,3.$$

تعريف : عدد حالات النجاح X في n من المحاولات لتجربة ذي الحدين يسمى متغير عشـــوائي يتبع ذي الحدين **binomial random variable** .

التوزيع الاحتمالي لمتغير عشوائي X يتبـــع ذي الحديـــن يســـمى توزيـــع ذي الحديـــن **binomial distribution** وسوف نرمز له بالرمز b(x;n,p) وذلك لأن قيمه تعتمـــد علـــى عدد المحاولات واحتمال النجاح في محاولة معطاه. وعلى ذلك عند إلقاء عملة 3 مـــرات فـــإن التوزيع الاحتمالي للمتغير العشوائي X سوف يعاد كتابته كالآتي :

$$b(x;3,\frac{1}{2}) = \frac{\binom{3}{x}}{8}, \quad x = 0,1,2,3.$$

الآن دعنا نعمم المثال السابق للحصول على صيغه لـ $b(x;n,p)$ • أي أننا نرغب في إيجاد صيغه تعطى احتمال الحصول على x حالات نجاح و $n-x$ حالات فشل في n من المحاولات لتجربة ذي الحدين • أولا نوجد احتمال الحصول على x حالات نجاح و $(n-x)$ حالات فشل في ترتيب معين • وحيث أن المحاولات مستقلة، فإنه يمكن ضرب كل الاحتمالات المقابلة للنواتج المختلفة • كل نجاح يحدث باحتمال p وكل فشل يحدث باحتمال $q=1-p$ • وعلى ذلك، الاحتمال للترتيب المعين هو $p^x q^{n-x}$ وبالطبع ترتيب الضرب هنا غير مهم • الآن لا بد من حساب العدد الكلي من نقاط العينة في التجربة (الترتيبات) والتي تعطى x حالات نجاح و $(n-x)$ حالات فشل • هذا العدد يساوى عدد الطرق لتبديل n من العناصر حيث x منهم من نوع (نجاح) و $n-x$ من نوع آخر (فشل) • وهذا يحدث بطرق عددها $\binom{n}{x}$ • وحيث أن الترتيبات التي عددها $\binom{n}{x}$ متنافية في إمكانية الحدوث وتحدث باحتمال متساوي ، فإنه يمكن الحصول على توزيع ذي الحدين بجمع الاحتمالات لكل الترتيبات أو ببساطة ضرب $p^x q^{n-x}$ في $\binom{n}{x}$ • وعلى ذلك فإن الصيغة العامة لتوزيع ذي الحدين سوف تكون على الشكل :

$$b(x;n,p)=\binom{n}{x}p^x q^{n-x},\ x=0,1,2,\dots n.$$

ويجب أن نتذكر أنه في مثال إلقاء عملة 3 مرات حيث $n=3, p=\frac{1}{2}$ فإن :

$$b(x;3,\frac{1}{2})=\binom{3}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x}=\frac{\binom{3}{x}}{8}$$

والتي تتفق مع الإجابة التي سبق أن حصلنا عليها •

في الحقيقة فإن التوزيع الاحتمالي لذي الحدين أشتق أسمه من أن $(n+1)$من الحدود في مفكوك ذي الحدين $(p+q)^n$ تقابل قيم $b(x;n,p)$ ، حيث $x=0,1,2,\dots n$ ، أي أن :

$$\begin{aligned}(p+q)^n &= \binom{n}{0}q^n+\binom{n}{1}pq^{n-1}+\binom{n}{2}p^2q^{n-2}+\dots+\binom{n}{n}p^n\\ &= b(0;n,p)+b(1;n,p)+\dots+b(n;n,p).\end{aligned}$$

وحيث أن $p+q=1$ فإننا نجد أن $\sum_{x=0}^{n} b(x;n,p)=1$ وهو الشرط الذي لابد من تحققه لأي توزيع إحتمالى • في الحقيقة سوف يكون الاهتمام بإيجاد $P(X\le r)$ أو $P(a\le X\le b)$ • لحسن الحظ فإن $\sum_{x=0}^{r} b(x;n,p)$ معطاة في جدول ذي الحدين في ملحق(١) حيث n=5,10,15,20,25.

مثال (٦-٣) إذا ألقيت زهرة نرد متزنة 6 مرات ، أوجد احتمال ظهور الرقم 5 أربع مرات بالضبط .

الحل . هنا يعتبر ظهور الرقم 5 نجاح وعلى ذلك احتمال النجاح في كل محاولة من المحاولات الستة المستقلة هو $\frac{1}{6}$. وعلى ذلك احتمال (الفشل) عدم ظهور الرقم خمسة هو $\frac{5}{6}$. بفرض أن X تمثل عدد مرات ظهور الرقم 5 ولها دالة كثافة الاحتمال :

$$b(x;n,p)=\binom{n}{x}p^{x}q^{n-x},\ x=0,1,2,\dots n.$$

وعلى ذلك الاحتمال المطلوب هو :

$$b(4;6,\frac{1}{6})=\binom{6}{4}(\frac{1}{6})^{4}(\frac{5}{6})^{2}$$

$$=\frac{6!}{4!2!}\cdot\frac{5^{2}}{6^{6}}=0.00803755.$$

مثال (٦-٤) صندوق به 10 ثمرات منها 3 تالفة، اختيرت منه ثمرتين . أحسب احتمال أن تكون واحدة تألفه (السحب بإرجاع) .

الحل . بفرض أن X تمثل عدد الثمار التالفة ولها دالة كثافة الاحتمال :

$$b(x;n,p)=\binom{n}{x}p^{x}q^{n-x},x=0,1,2,\dots n.$$

وعلى ذلك الاحتمال المطلوب هو :

$$b(1;2,0.3)=\binom{2}{1}(0.3)^{1}(0.7)^{1}$$

$$=\frac{2!}{1!1!}\cdot(0.3)^{1}(0.7)^{1}=0.42.$$

مثال (٦-٥) احتمال أن يشفى مريض من مرض نادر في الدم هو 0.2، فإذا كان معروف أن 15 شخص عندهم هذا المرض ما هو احتمال أن (ا) يشفى 9 على الأقل من المرض (ب) يشفى من 4 إلى 8 من هذا المرض (جـ) يشفى على الأكثر أثنين من هذا المرض؟

الحل . (ا) بفرض أن X تمثل عدد المرضى الذين سوف يشفوا من هذا المرض فإن :

$$P(x\geq 9)=1-P(x<9)=1-P(x\leq 8)$$

$$=1-\sum_{x=0}^{8}b(x;15,0.2)=1-0.999=0.001.$$

(ب)

$$P(4\leq X\leq 8)=\sum_{x=0}^{8}b(x;15,0.2)-\sum_{x=0}^{3}b(x;15,0.2)$$

$$=0.999-0.648=0.351.$$

(جـ)
$$P(X \le 2) = \sum_{x=0}^{2} b(x;15,0.2) = 0.3980.$$

نظرية (٦-١) الوسط الحسابي والتباين لتوزيع ذي الحدين $b(x;n,p)$ هما :

$$\mu = pq \ , \ \sigma^2 = npq$$

البرهان : بفرض أن نتيجة المحاولة رقم j يمثل بالمتغير العشوائي I_j والـذي يسـمى (متغـير برنولى) والذي يأخذ القيمة 0 أو 1 باحتمالات p,q على التوالي. أي أن $I_j = 0$ تدل علـى فشل و $I_j = 1$ تدل على نجاح. وعلى ذلك، في تجربة ذي الحدين يمكن كتابة عـدد حـالات النجاح كحاصل جمع لعدد n من متغيرات برنولى المستقلة. وعلى ذلك :

$$X = I_1 + I_2 + \ldots + I_n.$$

الوسط الحسابي لأي I_j هو $E(I_j) = 0 \cdot q + 1 \cdot p = p$ وعلى ذلك، وباستخدام نظريـة (٦-٤) فإن الوسط الحسابي لتوزيع ذي الحدين هو :

$$\mu = E(X) = E(I_1) + E(I_2) + \ldots + E(I_n)$$
$$= p + p + \ldots + p = np.$$

وذلك لأن عدد الحدود يساوى n .

التباين لأي I_j هو :

$$\sigma_{I_j}^2 = E(I_j - p)^2 = E(I_j^2) - p^2$$
$$= (0)^2 q + (1)^2 p - p^2 = p(1-p) = pq.$$

وعلى ذلك ، وباستخدام نظرية (٤-٨) فإن التباين لتوزيع ذي الحدين هو :

$$\sigma_X^2 = \sigma_{I_1}^2 + \sigma_{I_2}^2 + \ldots + \sigma_{I_n}^2$$
$$= pq + pq + \ldots + pq = npq.$$

مثال (٦-٦) أوجد الوسط الحسابي والتباين لتوزيع ذي الحدين في مثال (٦-٥) .

الحل . حيث $n = 15 \ , \ p = 0.2$ وباستخدام نظرية (٦-١) فإن :

$$\mu = (15)(0.2) = 3 \ , \sigma^2 = (15)(0.2)(0.8) = 2.4.$$

تجربة ذي الحدين تسمى تجربة متعددة الحـــدود **multinomial experiment** إذا كانت كل محاولة لها **k** من النواتج حيث k>2 . عموما إذا كانت لمحاولة معطاة **k** من النواتـج الممكنة $E_1, E_2, \ldots, E_k$ باحتمالات $p_1, p_2, \ldots, p_k$ ، فإن التوزيع المتعدد الحدود سـوف يعطى الاحتمال أن E_1 تحدث x_1 من المرات وأن E_2 تحدث x_2 من المـرات و...و E_k تحدث x_k من المرات في **n** من المحاولات المستقلة، حيث $x_1 + x_2 + \ldots + x_k = n$. سوف

نرمز للتوزيع المتعدد الحدود بالرمز $f(x_1, x_2, ..., x_k; p_1, p_2, ..., p_k, n)$ • من الواضح أن $p_1 + p_2 + ... + p_k = 1$ لأن نتيجة كل محاولة لابد أن تكون واحدة من k من النواتج الممكنة • لاشتقاق الصيغة العامة سوف نتبع نفس الأسلوب المستخدم في توزيع ذي الحدين • وحيث أن المحاولات مستقلة وأنه في ترتيب معين نحصل على x_1 نواتج من E_1 و x_2 من E_2 و • • • و x_k من E_k باحتمال $p_1^{x_1} p_2^{x_2} ... p_k^{x_k}$ • العدد الكلى من الترتيبات والذي يؤدى إلى نفس النتائج في n من المحاولات المستقلة يساوى عدد الطرق لتبديل n من العناصر حيث x_1 منها تنتمي إلى E_1 و x_2 تنتمي إلى E_2 و • • • و x_k تنتمي إلى E_k والذي يحدث بطرق عددها $\frac{n!}{x_1! x_2! ... x_k!}$ • وحيث أن الترتيبات (نقاط العينة) التي عددها $\frac{n!}{x_1! x_2! ... x_k!}$ متنافية في إمكانية الحدوث وتحدث باحتمال متساوي ، فإنه يمكن الحصول على التوزيع المتعدد الحدود بضرب الاحتمال لترتيب معين في $\frac{n!}{x_1! x_2! ... x_k!}$ •وعلى ذلك فإن التوزيع متعدد الحدود يكون على الصورة :

$$f(x_1, x_2, ..., x_k; p_1, p_2, ..., p_k, n)$$
$$= \frac{n!}{x_1! x_2! ... x_k!} p_1^{x_1} p_2^{x_2} ... p_k^{x_k},$$

حيث $\sum_{i=1}^{k} x_i = n$, $\sum_{i=1}^{k} p_i = 1$ • وقد اشتق التوزيع المتعدد الحدود اسمه من أن الحدود في المفكوك $(p_1 + p_2 + ... + p_k)^n$ تقابل كل القيم الممكنة للدالة $f(x_1, x_2, ..., x_k; p_1, p_2, ..., p_k, n)$ •

مثال (٦-٧) ألقيت زهرة نرد متزنة 11 مرة أوجد احتمال الحصول على رقم واحد مرة ، ورقم 2 مرتين ورقم 3 خمس مرات ورقم أربعة مرتين ورقم 6 مرة واحدة •

الحل • احتمال ظهور أي رقم عند إلقاء زهرة نرد متزنة مرة واحدة متساوي ويساوى $\frac{1}{6}$ وعلى ذلك فإن الاحتمال المطلوب هو :

$$f(1,2,5,2,1;\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},11)$$
$$=\frac{11!}{1!2!5!2!1!}(\frac{1}{6})^1(\frac{1}{6})^2(\frac{1}{6})^5(\frac{1}{6})^2(\frac{1}{6})^1$$
$$=\frac{11!}{1!2!5!2!1!}\cdot\frac{1}{6^{11}}=0.000229219.$$

(٦-٣) التوزيع الهندسي الزائدى **Hypergeometric Distribution**

بفرض أن مجتمع يتكون من عدد محدود من الوحدات، وليكن N ، وأن هناك k من الوحدات من النوع A (نجاح) والوحدات الباقية من نوع B (فشل)، و بفرض أن عينة عشوائية من الحجم n اختيرت من هذا المجتمع بدون إرجاع. بفرض أن X تمثل عدد الوحدات من نوع A التي تظهر في العينة. اهتمامنا سوف يكون في إيجاد P(X=x). التجربة السابقة تسمى تجربة الهندسي الزائدى hypergeometric experiment . تحقق تجربة الهندسي الزائدى الشروط التالية :

-ا- عينة عشوائية من الحجم n تختار من مجتمع يحتوى على N من الوحدات.

-ب- في المجتمع الذي حجمه N فإن k من الوحدات تصنف نجاح و $N-k$ تصنف فشل.

تعريف : عدد حالات النجاح في تجربة الهندسي الزائدى يسمى متغير عشوائي يتبع الهندسي الزائدى.

التوزيع الاحتمالي لمتغير عشوائي يتبع التوزيع الهندسي الزائدى يسمى التوزيع الهندسي الزائدى ويمثل بالرمز $h(x;N,n,k)$ وذلك لأن عدد حالات النجاح x تعتمد على k الموجودة في الفئة N ،حيث يختار من N وحدات عددها n . في المثال (٦-٤) استخدمنا توزيع ذي الحدين إذا كان السحب بإرجاع (المحاولات مستقلة) . الآن بفرض أن السحب بدون إرجاع (المحاولات غير مستقلة) . في هذه الحالة سوف يكون هناك $\binom{3}{1}$ طريقة لاختيار ثمرة تالفة ولكل واحدة من هذه الطرق يوجد $\binom{7}{1}$ طريقة لاختيار ثمرة سليمة. وعلى ذلك عند اختيار ثمرتين من الصندوق بدون إرجاع فإن عدد الطرق الكلية للحصول على ثمرة تالفة وثمرة سليمة هو $\binom{3}{1}\binom{7}{1}$. العدد الكلى من الطرق لاختيار ثمرتين من الصندوق المحتوى على 10 ثمرات هو $\binom{10}{2}$.وعلى ذلك احتمال الحصول على ثمرة تالفة وثمرة سليمة عند اختيار عينة من الحجم n=2 من الصندوق المحتوى على 10 ثمرات هو :

$$\frac{\binom{3}{1}\binom{7}{1}}{\binom{10}{2}}=\frac{7}{15}.$$

المثال السابق يوضح ما يسمى بتجربة الهندسي الزائدى.

مثال (٦-٨) يراد اختيار لجنة من ثلاثة أشخاص من بين 4 سيدات و 5 رجال والمطلوب إيجاد التوزيع الاحتمالي لعدد السيدات في اللجنة المختارة.

الحل . بفرض أن المتغير العشوائي X يمثل عدد السيدات في اللجنة المختارة. الشروط لتجربة الهندسي الزائد متوفرة وعلى ذلك :

$$P(X=0)=h(0;9,3,4)=\frac{\binom{4}{0}\binom{5}{3}}{\binom{9}{3}}=\frac{10}{84},$$

$$P(X=1)=h(1;9,3,4)=\frac{\binom{4}{1}\binom{5}{2}}{\binom{9}{3}}=\frac{40}{84},$$

$$P(X=2)=h(2;9,3,4)=\frac{\binom{4}{2}\binom{5}{1}}{\binom{9}{3}}=\frac{30}{84},$$

$$P(X=3)=h(3;9,3,4)=\frac{\binom{4}{3}\binom{5}{0}}{\binom{9}{3}}=\frac{4}{84}.$$

ويمكن تمثيل التوزيع الهندسي الزائدى بالجدول التالي :

x	0	1	2	3
P(X=x)	$\frac{10}{84}$	$\frac{40}{84}$	$\frac{30}{84}$	$\frac{4}{84}$

هذا ويمكن وضع صيغة لإيجاد التوزيع الاحتمالي للمثال السابق على الشكل :

$$P(X=x)=h(x;9,3,4)=\frac{\binom{4}{x}\binom{5}{3-x}}{\binom{9}{3}},\ x=0,1,2,3.$$

الآن يمكن تعميم صيغة التوزيع الاحتمالي في المثال السابق وذلك للحصول على صيغة للدالة $h(x;N,n,k)$. العدد الكلى للعينات من الحجم n المختارة من N من الوحدات هو $\binom{N}{n}$. هذه العينات يفترض أنها متساوية في إمكانية الحدوث. يوجد $\binom{k}{x}$ طريقة لاختيار x

من k حالات النجاح ، ولكل طريقة من هذه الطرق يمكن اختيار (n-x) من حالات الفشل بطرق عددها $\binom{N-k}{n-x}$. وعلى ذلك العدد الكلى من العينات المرغوب فيها من $\binom{N}{n}$ عينة هو $\binom{k}{x}\binom{N-k}{n-x}$. وعلى ذلك فإن التوزيع الهندسي الزائدى يمكن الحصول عليه كالآتي :

$$h(x;N,n,k)=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, x=0,1,2,...,n.$$

مثال (٦-٩) اختيرت عينة عشوائية من الحجم n=6 من صندوق يحتوى على 5 كرات حمراء و 4 كرات سوداء . ما هو احتمال ظهور ثلاث كرات حمراء في العينة المختارة .

الحل . باستخدام التوزيع الهندسي الزائدى حيث $x=3, k=5, n=6, N=9$ ، فإن :

$$h(3;9,6,5)=\frac{\binom{5}{3}\binom{4}{3}}{\binom{9}{6}}=\frac{40}{84}.$$

نظرية (٦-٢) الوسط الحسابي والتباين للتوزيع الهندسي الزائدى هما :

$$\mu=\frac{nk}{N}, \sigma^2=\frac{N-n}{N-1}\cdot n\cdot\frac{k}{N}(1-\frac{k}{N}).$$

إذا كانت n صغيرة بالنسبة إلى N فإنه يمكن استخدام توزيع ذي الحدين كتقريب للتوزيع الهندسي الزائدى حيث $p=\frac{k}{N}$.وعلى ذلك يمكن تقريب الوسط الحسابي والتباين بالصيغة :

$$\mu=np=\frac{nk}{N}, \sigma^2=npq=n\cdot\frac{k}{N}(1-\frac{k}{N}).$$

بمقارنة الصيغتين السابقتين بالصيغتين في نظرية (٦-٢) فإننا نجد أن الوسط الحسابي هو نفسه بينما التباين يختلف بمعامل تصحيح $\frac{N-n}{N-1}$ وهذا يمكن إهماله عندما تكون n صغيرة بالنسبة إلى بالنسبة N .

مثال (٦-١٠) في سنترال وجد أنه من بين 4000 تليفون تم تركيبهم في منطقة حديثة يوجد 3000 منهم يختلف لونهم عن اللون الأسود . تحدث 5 أشخاص عشوائيا، فما هو احتمال أن 2 منهم بالضبط سوف يتحدثون من تليفون لونه أسود .

الحل . وحيث أن المجتمع حجمه N=4000 وهذا الحجم كبير نسبيا بالنسبة لحجم العينة n=5 فإننا سوف نقرب الاحتمال المطلوب باستخدام توزيع ذي الحدين . احتمال أن شخص يتحدث من تليفون أسود هو 0.25 وعلى ذلك احتمال أن 2 بالضبط يتصلون من تليفون أسود هو :

$$h(2;4000,5,1000) \approx b(2;5,0.25)$$
$$= \sum_{x=0}^{2} b(x;5,0.25) - \sum_{x=0}^{1} b(x;5,0.25) = 0.896 - 0.633 = 0.263.$$

يمكن تعميم التوزيع الهندسي الزائدى ليعامل للحالة التي يكون فيها المجتمع مقسم إلى k من الخلايا $A_1, A_2, ..., A_k$ حيث a_1 وحدة في الخلية A_1 و a_2 وحدة في الخلية A_2 و ... و a_k وحدة في الخلية A_k . الآن سوف يكون اهتمامنا في إيجاد صيغة لاحتمال أن العينة العشوائية من الحجم n سوف تعطى x_1 عنصر من A_1 و x_2 عنصر من A_2 و ... و x_k عنصر من A_k والتي سوف تمثل الاحتمال بالصيغة :

$$f(x_1, x_2, ..., x_k; a_1, a_2, ..., a_k, N, n).$$

لإيجاد صيغة عامة فإننا نعرف أن العدد الكلى من العينات من الحجم n التي يمكن اختيارها من مجتمع حجمه N هو $\binom{N}{n}$. وعلى ذلك هناك $\binom{a_1}{x_1}$ طريقة لاختيار x_1 من الوحدات الموجودة في A_1 ولكل واحدة من هذه الطرق فإننا يمكن اختيار x_2 من الوحدات الموجودة في A_2 بطرق عددها $\binom{a_2}{x_2}$. وعلى ذلك يمكن اختيار x_1 من الوحدات الموجودة في A_1 و اختيار x_2 من الوحدات الموجودة في A_2 بطرق عددها $\binom{a_1}{x_1}\binom{a_2}{x_2}$. وبالاستمرار في هذا الأسلوب يمكن اختيار كل الوحدات التي عددها n والتي تحتوى على x_1 من الوحدات الموجودة في A_1 و x_2 من الوحدات الموجودة في A_2 و ... و x_k من الوحدات الموجودة في A_k بطرق عددها $\binom{a_1}{x_1}\binom{a_2}{x_2}..\binom{a_k}{x_k}$. وعلى ذلك فإن التوزيع الاحتمالي المطلوب سوف يكون :

$$f(x_1, x_2, ..., x_k; a_1, a_2, ..., a_k, N, n) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2}..\binom{a_k}{x_k}}{\binom{N}{n}}.$$

حيث $\sum_{i=1}^{k} a_i = N, \quad \sum_{i=1}^{k} x_i = n$.

مثال (٦-١١) يحتوى صندوق على 3 كرات حمراء و 4 كرات بيضاء و 5 كرات سوداء. يرغب شخص في اختيار 4 كرات من الصندوق، ما هو الاحتمال أن يختار كرة حمراء و2 كرة بيضاء وواحدة سوداء ؟

الحل . باستخدام الصيغة العامة للتوزيع الهندسي الزائدى حيث $a_1 = 3, x_1 = 1, a_2 = 4, x_2 = 2, a_3 = 5, x_3 = 1, N = 12, n = 4$ فإن :

$$f(1,2,1;3,4,5,12,4) = \frac{\binom{3}{1}\binom{4}{2}\binom{5}{1}}{\binom{12}{4}} = \frac{90}{495}.$$

(٦-٤) توزيع بواسون **Poisson Distribution**

أن التجارب التي تعطينا عدد حالات النجاح والتي تحدث في فترة زمنية معينة أو في منطقة محددة، تسمى تجارب بواسون poisson experiment . الفترة الزمنية المعينة قد تكون دقيقة، يوم، أسبوع، شهر أو حتى سنة. وعلى ذلك تجربة بواسون قد تنتج مشاهدات لمتغير عشوائي X يمثل عدد المكالمات التليفونية في الساعة والمستقبلة من مكتب، أو عدد الأيام في السنة والتي تغلق فيها بعض المدارس بسبب الصقيع في بلد ما. المنطقة المحددة يمكن أن تكون خط الأعداد الحقيقية، مساحة، حجم أو ربما قطعة من المعدن. في هذه الحالة X يمكن أن تمثل عدد الفئران في فدان من القمح، عدد البكتريا في لتر من الماء النقي، عدد الأخطاء في صفحة من قاموس. تجارب بواسون يجب أن تحقق الشروط التالية :

أ- متوسط عدد حالات النجاح، μ ، والتي تحدث في الفترة الزمنية المعطاة أو في المنطقة المحددة معلوم.

ب- احتمال وقوع حالة نجاح واحدة في فترة قصيرة أو منطقة صغيرة يتناسب مع طول هذه الفترة أو حجم هذه المنطقة ولا يعتمد على عدد حالات النجاح التي تحدث خارج هذه الفترة أو المنطقة.

جـ- احتمال وقوع أكثر من حالة نجاح في الفترة القصيرة أو المنطقة الصغيرة مهمل.

تعريف : عدد حالات النجاح X في تجربة بواسون يسمى متغير عشوائي يتبع بواسون.

التوزيع الاحتمالي لمتغير عشوائي X يتبع توزيع بواسون يمثل بالصيغة $p(x;\mu)$ وذلك لأن قيمته تعتمد فقط على μ . متوسط عدد حالات النجاح التي تحدث في الفترة المعينة أو المنطقة المحددة تساوى التباين. اشتقاق توزيع بواسون خارج نطاق هذا الكتاب. التوزيع الاحتمالي لمتغير عشوائي يتبع بواسون هو :

$$p(x;\mu) = \frac{e^{-\mu}\mu^{x}}{x!}, x = 0,1,2,\ldots,$$

حيث μ متوسط عدد حالات النجاح التي تحدث في الفترة المعطاة أو المنطقة الخاصة و **e=2.71828...** .

الجدول في ملحق (٢) يحتوى على حاصل جمع احتمالات بواسون أي $\sum_{x=0}^{r} p(x;\mu)$ لقيم محددة من μ تتراوح بين 0.1 إلى 20 . سوف نشرح طريقة عمل هذه الجداول بالمثالين التالين.

مثال (٦-١٢) تعطُل ماكينة لتصنيع الحلوى في المتوسط خمس مرات في الأسبوع ما هو الاحتمال أن تعطل الماكينة ثلاث مرات خلال أسبوع.

الحل . باستخدام توزيع بواسون حيث $\mu = 5,\ x = 3$ ومن جدول بواسون في ملحق (٢) فإن :

$$p(3;5) = \frac{e^{-5}5^3}{3!} = \sum_{x=0}^{3} p(x;5) - \sum_{x=0}^{2} p(x;5)$$

$$0.265 - 0.125 = 0.14.$$

مثال (٦-١٣) إذا كان متوسط عدد الفئران في فدان من القمح هو $\mu = 3$. أوجد احتمال وجود على الأقل 12 فأرا في فدان معطى.

الحل . بفرض أن X تمثل عدد الفئران في فدان من القمح وباستخدام جداول بواسون في ملحق (٢) فإن :

$$P(X \geq 12) = 1 - P(X < 12)$$

$$== 1 - \sum_{x=0}^{11} p(x;3)$$

$$= 1 - P(X \leq 11) = 1 - 1.000 = 000.0.$$

المدرج التكراري لكل من توزيع ذي الحدين وتوزيع بواسون تقريبا لهم نفس الشكل عندما تكون n كبيرة و p تقترب من الصفر وعلى ذلك إذا تحقق هذا الشرط، فإن توزيع بواسون بمتوسط $\mu = np$، يمكن استخدامه كتقريب لإيجاد احتمالات توزيع ذي الحدين. إذا اقتربت p من الواحد يمكن تغيير ما قمنا بتعريفه نجاح إلى فشل والعكس الفشل إلى نجاح . وعلى ذلك تتغير p إلى قيمة قريبة من الصفر

مثال (٦-١٤) يحتوى كتاب على 1000 صفحة يوجد بها 400 خطا موزعة على صفحات الكتاب ما هو احتمال الحصول على ستة أخطاء في صفحة مختارة ؟

الحل . المتغير العشوائي X يمثل عدد الأخطاء في إحدى الصفحات وهو يتبع توزيع ذي الحدين بمعلمتين n , p حيث $p = \frac{1}{1000}$ و n=400 وحيث أن n كبيرة و p صغيرة فإن توزيع X يؤول إلى توزيع بواسون بمعلمة $\mu = np = (400)(\frac{1}{1000}) = 0.4$ و باستخدام جداول بواسون في الملحق (٢) حيث $\mu = 0.4$, فإن :

$$p(x = 6) = b(x;400,0.001) \approx p(6;0.4),$$

$$p(6;0.4)=\frac{e^{-0.4}0.4^{6}}{6!}=\sum_{x=0}^{6}p(x;0.4)-\sum_{x=0}^{5}p(x;0.4)=1.000-1.000=000.0.$$

مثال (٦-١٥) بفرض أنه في المتوسط يوجد شخص من بين 1000 شخص يدخن السجائر في مدينة ما . أوجد احتمال أنه في عينة عشوائية من 8000 شخص نحصل على 6 أشخاص يدخنون .

الحل . في هذا المثال يكون لدينا تجربة ذي الحدين حيث n=8000 ، p=0.001 وحيث أن p تقترب كثيرا من الصفر وأن n كبيرة بدرجة كافية، فإنه يمكن استخدام توزيع بواسون حيث $\mu=(8000)(0.001)=8$ **. وعلى ذلك إذا كانت X تمثل عدد المدخنين، فإن :**

$$\begin{aligned}p(X=6)&=b(x;8000,0.001)\\&\approx p(6;8)=\sum_{x=0}^{6}p(x;8)-\sum_{x=0}^{5}p(x;8)\\&=0.313-0.191=0.122.\end{aligned}$$

(٦-٥) توزيع ذي الحدين السالب Negative Binomial Distribution

بفرض أن تجربة ما لها نفس الخصائص التي سبق أن ذكرناها لتوزيع ذي الحدين، ولكن مع تكرار المحاولات حتى يمكن الحصول على عدد ثابت من حالات النجاح. وعلى ذلك، بدلا من إيجاد احتمال الحصول على x نجاح في n من المحاولات، فإن الاهتمام سوف يكون في إيجاد احتمال أن النجاح رقم k سوف يحدث في المحاولة رقم x . التجارب من هذا النوع تسمى تجارب ذي الحدين السالب negative binomial experiment . بفرض أن لاعب كرة السلة ينجح في التصويب نحو الهدف في 80% من المحاولات . المطلوب إيجاد احتمال أن ينجح في تصويب الهدف الخامس في المحاولة رقم 8 . سوف نرمز للنجاح في التصويب بالرمز F′ ونرمز للفشل في التصويب بالرمز F ، وعلى ذلك لترتيب مطلوب سوف نحصل على F′F′FFFF′F′F′ والذي يحدث باحتمال: $(0.8)(0.8)(0.2)(0.2)(0.2)(0.8)(0.8)(0.8)=(0.8)^5(0.2)^3$ **. ويمكن حصر كل الترتيبات بإعادة ترتيب حالات النجاح والفشل ما عدا المحاولة الأخيرة، والتي لابد أن تكون النجاح رقم خمسة. العدد الكلى من الترتيبات الممكنة يساوى عدد الطرق لتبديل 7 عناصر منها 4 من نوع نجاح و 3 من نوع فشل. هذا العدد الكلى من الترتيبات الممكنة يحدث بطرق متنافية عددها** $\binom{7}{4}$ **. وحيث أن X تمثل عدد المحاولات اللازمة للحصول على 5 أهداف، فإن :**

$$P(X=8)=\binom{7}{4}(0.8)^5(0.2)^3=0.0917504.$$

تعريف : العدد X من المحاولات والذي ينتج k حالات نجاح في تجربة ذي الحدين السالب يسمى متغير عشوائي يتبع ذي الحدين السالب .

التوزيع الاحتمالي لمتغير عشوائي يتبع ذي الحدين السالب يسمى توزيـــع ذي الحديـــن السالب، وسوف نمثله بالصيغة $b^*(x;k,p)$ حيث أن قيمه تعتمد على عدد حالات النجـــاح المطلوبة واحتمال النجاح في المحاولة المعطاه . لإيجاد الصيغة العامة سوف نوجـــد أولا احتمـــال الحصول على

نجاح في المحاولة رقم x والتي تكون مسبوقة بعدد (k-1) من حالات النجاح و (x-k) من حالات الفشل وذلك في ترتيب خاص . وحيث أن المحاولات مستقلة، فإنه يمكن ضرب كل الاحتمالات المقابلة لكل نتيجة مطلوبه . كل حالة نجاح تحدث باحتمال p وكل حالة فشل تحدث باحتمـــال q=1-p . وعلـــى ذلـــك، الاحتمـــال أن الـــترتيب المطلـــوب ينتـــهي بحالـــة نجـــاح هـــو $p^{k-1}q^{x-k}p = p^kq^{x-k}$. العدد الكلى من نقاط العينة في التجربة والذي ينتـــهي بحالـــة نجاح، بعد k-1 من حالات النجاح و x-k من حالات الفشل في أي ترتيـــب، يســـاوى عـــدد الطرق لتبديل x-1 من المحاولات منهم k-1 حالات نجـــاح و x-k حـــالات فشـــل والـــتي تساوى $\binom{x-1}{k-1}$. كل الترتيبات المطلوبة تكون متنافية في إمكانية الحدوث وتحـــدث باحتمـــال متساوي وهـــو p^kq^{n-k}، وعلـــى ذلـــك يمكـــن الحصـــول علـــى الاحتمـــال المطلـــوب بضرب p^kq^{x-k} في $\binom{x-1}{k-1}$، وعلى ذلك توزيع ذي الحدين السالب يأخذ الصيغة التالية :

$$b^*(x;k,p) = \binom{x-1}{k-1}p^kq^{x-k}, \quad x = k, k+1, k+2, \ldots$$

مثال (٦-١٦) إذا كان احتمال ولادة ذكر في أي ولادة تمر بها سيدة هو $\frac{1}{2}$ أوجد احتمال أن تضع ذكرين بعد أربع ولادات .

الحل . باستخدام توزيع ذي الحدين السالب حيث $x = 4,\ k = 2,\ p = \frac{1}{2}$ فإن :

$$b^*(4;2,0.5) = \binom{3}{1}0.5^2 0.5^2$$

$$= \frac{3!}{1!2!} \cdot (\frac{1}{2})^2 (\frac{1}{2})^2 = \frac{3}{16}.$$

في الحقيقة أشتق اسم توزيع ذي الحدين السالب من أن كل حد في المفكوك $p^k(1-q)^{-k}$ يقابل قيمة من قيم $b^*(x;k,p)$ حيث $x = k, k+1, k+2, \ldots$ •
عندما n=1 فإننا نحصل على الحالة الخاصة من توزيع ذي الحدين السالب، أي نحصل على التوزيع الاحتمالي لعدد المحاولات المطلوبة للحصول على حالة نجاح واحدة • توزيع ذي الحدين السالب سوف يختزل إلى الشكل :

$$b^*(x;1,p) = pq^{x-1}, x = 1,2,3,\ldots$$

والذي يسمى التوزيع الهندسي وسوف نرمز له بالرمز $g(x;p)$ •

مثال (٦-١٧) أوجد (أ) التوزيع الاحتمالي لعدد المحاولات اللازمة للحصول على صورة واحدة وذلك عند إلقاء عملة متزنة (ب) احتمال الحصول على صورة في المحاولة الرابعة •

الحل • باستخدام التوزيع الهندسي حيث $p = \frac{1}{2}$ نحصل على :

(أ)
$$g(x;0.5) = (0.5)(0.5)^{x-1}, x = 1,2,3,\ldots$$

(ب)
$$g(4;\frac{1}{2}) = (0.5)(0.5)^3 = \frac{1}{16}.$$

(٦-٦) التوزيع الطبيعي Normal Distribution

يعتبر التوزيع الطبيعي من أهم التوزيعات الاحتمالية المتصلة في علم الإحصاء، حيث يصف كثيرا من المجتمعات الموجودة في الطبيعة ، الصناعة ، الأبحاث. دالة كثافة الاحتمال للتوزيع الطبيعي هي:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

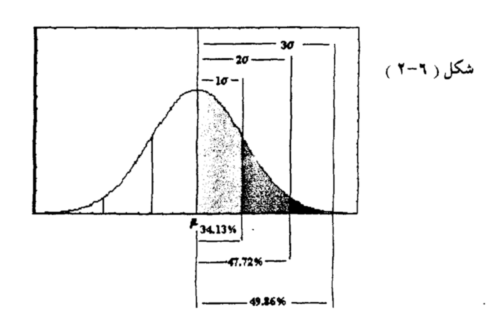
على الفترة $-\infty < x < \infty$ حيث $e=2.71828\ldots$ و $\pi=3.14159\ldots$ و μ و σ هما الوسط الحسابي (المتوسط) والانحراف المعياري على التوالي. بيان $f(x)$ في شكل (٦-٢) على شكل ناقوس متماثل حول العمود المقام على الوسط الحسابي • ينطبق الوسط الحسابي على الوسيط وأيضا على المنوال • يتقارب طرفا منحنى التوزيع الطبيعي من الصفر عندما $x \to \infty$ أو $x \to -\infty$ •

لأي توزيع طبيعي فإن :-

(ا) المساحة الواقعة تحت المنحني الطبيعي بين النقطتين μ و $(\mu + \sigma)$ تمثل 34.13% من المساحة الكلية كما يتضح من شكل (٦-٢).

(ب) المساحة الواقعة تحت المنحنى الطبيعي بين النقطتين μ و $(\mu+2\sigma)$ تمثل **47.72** من المساحة الكلية كما يتضح من شكل (٦-٢).

(جـ) المساحة الواقعة تحت المنحنى الطبيعي بين النقطتين μ و $(\mu+3\sigma)$ تمثل **% 49.86** من المساحة الكلية كما يتضح من شكل (٦-٢).



شكل (٦-٢)

وبما أن المنحنى الطبيعي متماثل فإن القيم السابقة تتحقق عند طرح الانحراف المعياري من المتوسط.

(٦-٦-١) التوزيع الطبيعي القياسي :

Standard Normal Distribution

إذا كان التوزيع الطبيعي متوسطه صفر وتباينه الواحد الصحيح فإنه يسمي التوزيع الطبيعي القياسي . بفرض أن Z ترمز لمتغير عشوائي متصل له توزيع طبيعي قياسي فإن دالة الكثافة الاحتمالية لذلك المتغير تأخذ الشكل الآتي :

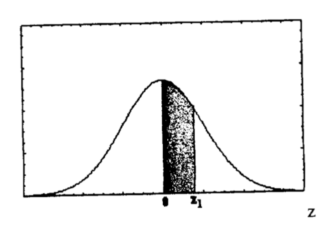
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad \infty \le z \le \infty.$$

إذا كان z_1 عدد حقيقي موجب فان الاحتمال $(0 < Z < z_1)$ يساوي المساحة المظللة في شكل (٦-٣)، ويمكن الحصول عليها من جدول التوزيع الطبيعي القياسي في ملحق (٣). المساحات الواقعة تحت المنحني الطبيعي القياسي غير معطاة في الجدول لقيم z السالبة ولكن يمكن حسابهم باستخدام خاصية التماثل للمنحنى الطبيعي. المساحة الواقعة تحت المنحنى الطبيعي

القياسي بين $z=0$ و $z=-z_1$ تســاوي المساحة الواقعة بين $z=z_1$ و $z=0$ أي أن :

$$P(-z_1 < Z < 0) = P(0 < Z < z_1)$$

معظم المساحات تحت المنحنى الطبيعي القياسي تقع في الفترة (3,3-) ونادرا ما نجد قيم تقع خارج هذه الفترة .



شكل (٦-٣)

مثال (٦-١٨) إذا كان **Z** متغير عشوائي يتبع التوزيع الطبيعي القياسي احسب الاحتمالات الآتية مع توضيح ذلك بيانيا .

(أ) $P(0 < Z < 1.05)$ (ب) $P(-1.06 < Z < 1.06)$

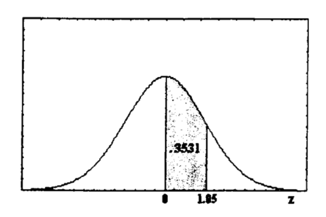
(جـ) $P(-0.47 < Z < 0.95)$ (د) $P(1.6 < Z < 2)$

(هـ) $P(Z > 2.02)$ (و) $P(Z < -0.45)$ (ز) $P(Z < 1.07)$

الحل . (أ) لإيجاد قيمة الاحتمال $P(0 < Z < 1.05)$ نبحث في العمود الأول على الشمال من جدول التوزيع الطبيعي القياسي في ملحق (٣) عن القيمة **1.0** ثم نتحرك أمام هذه القيمة أفقيا حتى نصل إلى العمود الذي رأس عنوانه الرقم **0.05** فتكون هي المساحة المطلوبة أي أن :

$$P(0 < Z < 1.05) = 0.3531.$$

والتي تمثل المساحة المظللة في شكل (٦-٤) .



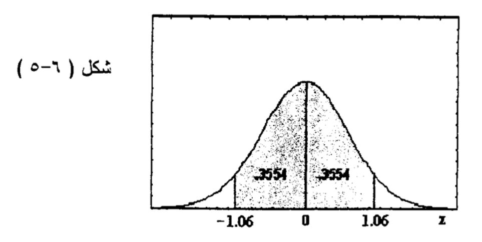
شكل (٦-٤)

(ب) الاحتمال المطلوب هو $P(-1.06 < Z < 1.06)$ وهو يساوي المساحة المظللة في شـكل (٦-٥) •

ونظرا لتماثل المنحنى الطبيعي فان :

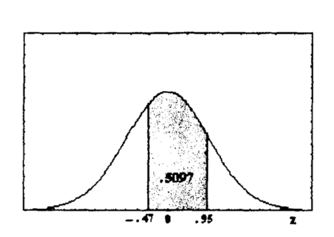
$$P(-1.06 < Z < 1.06) = P(-1.06 < Z < 0) + P(0 < Z < 1.06)$$
$$= 2P(0 < Z < 1.06) = 2(0.3554) = 0.7108.$$

شكل (٦-٥)



(جـ) الاحتمال المطلوب هو $P(-0.47 < Z < 0.95)$ وهو يساوي المســاحة المظللــة في شكل (٦-٦) • ونظرا لتماثل المنحنى الطبيعي فإن :

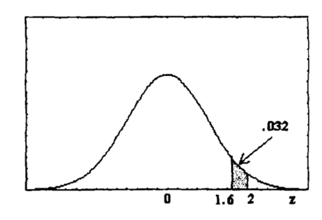
$$P(-0.47 < Z < 0.95) = P(-0.47 < Z < 0) + P(0 < Z < 0.95)$$
$$= P(0 < Z < 0.47) + P(0 < Z < 0.95) = 0.1808 + 0.3289 = 0.5097.$$



شكل (٦-٦)

(د) الاحتمال المطلوب هو $P(1.6 < Z < 2)$ وهو يساوي المساحة المظللة في شكل(٦-٧) .
أي أن :

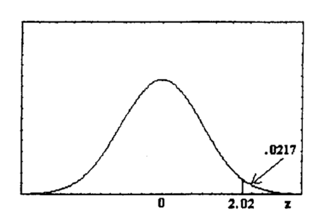
$$P(1.6 < Z < 2) = P(0 < Z < 2) - P(0 < Z < 1.6)$$
$$= 0.4772 - 0.4452 = 0.032.$$



شكل (٦-٧)

(هـ) من شكل (٦-٨) وباستخدام حقيقة أن $P(Z > 0)$ يساوي نصف المساحة الكلية تحت المنحنى الطبيعي أي أن $P(Z > 0) = 0.5$ وعلى ذلك :

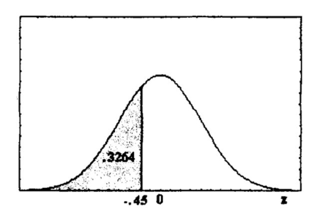
$$P(Z > 2.02) = P(Z > 0) - P(0 < Z < 2.02)$$
$$= 0.5 - 0.4783 = 0.0217.$$



شكل (٦-٨)

(و) الاحتمال المطلوب هو $P(Z < -0.45)$ وهو يساوي المساحة المظللة في شـــكل (٦-٩)
ونظرا لتماثل المنحنى الطبيعي فإن :

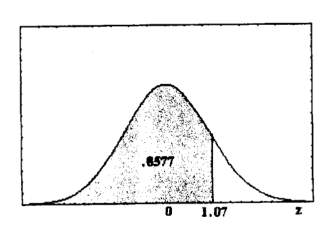
$$P(Z < -0.45) = P(Z > 0.45)$$
$$= P(Z > 0) - P(0 < Z < 0.45)$$
$$= 0.5 - 0.1736 = 0.3264.$$



شكل (٦-٩)

(ز) الاحتمال المطلوب هو $P(Z < 1.07)$ وهو يساوي المساحة المظللة في شـــكل (٦-١٠).
ولان المنحنى الطبيعي متماثل ومساحة كل جانب من جانبي المنحنى تساوي **0.5** فان :

$$P(Z<1.07)=P(Z<0)+P(0<Z<1.07)$$
$$=0.5+0.3577=0.8577.$$



شكل (٦-١٠)

<u>استخدام التوزيع الطبيعي القياسي لاستخراج المساحات تحت المنحنى الطبيعي</u>

الآن نعود مرة أخري إلى حالة متغير عشوائي طبيعـــي متوســـطه μ وانحرافـــه المعياري σ . يمكن تحويل المتغير X إلى متغير طبيعي قياسي Z باستخدام الصيغة التالية :

$$Z=\frac{X-\mu}{\sigma}.$$

التحويل من X إلى Z يمثل انتقال لنقطة الأصل مصحوبا بتغير لمقياس الرسم. عندمـــا $x=\mu$ فإن $z=0$ ، وعندما $x=\mu-\sigma$ فإن $z=-1$ ، وعندما $x=\mu+2\sigma$ فإن $z=2$ وهكذا. أي أن مقياس الرسم قد تغير حيث تناظر مسافة σ على محور السينات مسافة قدرها واحد على محور z ، وعلى ذلك يمكن استخدام جداول التوزيع الطبيعي القياسي في ملحق (٣) لحساب الاحتمالات لأي متغير عشوائي طبيعي كما يتضح من الأمثلة التالية .

مثال (٦-١٩) في مدينة صغيرة وجد أن أعلى درجة حرارة مسجلة يوميا خلال فصل الربيع لها متوسط 20°c وانحراف معياري 5°c. بفرض أن المتغير العشوائي X (أعلى درجة حـــرارة يوميا) يخضع للتوزيع الطبيعي، أوجد النسبة المئوية للأيام التي فيها أعلى درجة حرارة :

(أ) بين 22°c و 26°c

(ب) على الأقل 28°c

الحل . (أ) إذا كان X يرمز لأعلى درجة حرارة مسجلة يوميا فان X يكون متغـــير عشـــوائي طبيعي متوسطه $\mu=20$ وانحرافه المعياري $\sigma=5$. المتغير الطبيعي القياسي المناظر هو :

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 20}{5}.$$

عندما $x_1 = 22$ فإن :

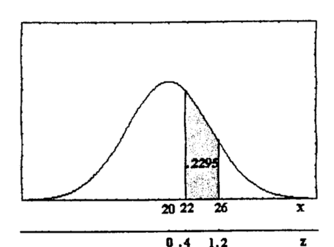
$$z_1 = \frac{22 - 20}{5} = 0.4.$$

وعندما $x_2 = 26$ فان :

$$z_2 = \frac{26 - 20}{5} = 1.2.$$

الاحتمال المطلوب هو $P(22<X<26)$ وهو يساوي المساحة المظللة في شكل (٦-١١) .
أي أن :

$$P(22 < X < 26) = P(0.4 < Z < 1.2)$$
$$= 0.3849 - 0.1554 = 0.2295.$$



شكل (٦-١١)

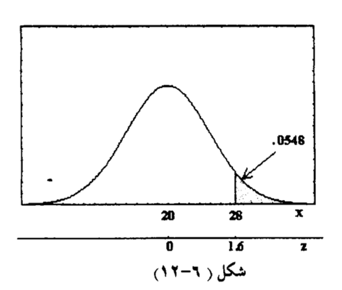
أي أن النسبة المئوية للأيام التي فيها أعلى درجة حرارة بين (22°c و 26°c) هي % 22.95 .

(ب) عندما $x_1 = 28$ فإن :

$$z_1 = \frac{28 - 20}{5} = 1.6.$$

الاحتمال المطلوب هو $P(X \geq 28)$ وهو يساوي المساحة المظللة في شكل (٦-١٢) . أي أن

$$P(X > 28) = P(Z > 1.6)$$
$$= P(Z > 0) - P(0 < Z < 1.6)$$
$$= 0.5 - 0.4452 = 0.0548.$$



شكل (٦-١٢)

أى أن النسبة المئوية للأيام التي فيها أعلى درجة حرارة فوق 28°c هي 5.48 % .

مثال (٦-٢٠) إذا كانت مبيعات الغاز الطبيعي في الأسبوع في محطة لتعبئة الغاز يخضع للتوزيع الطبيعي بمتوسط $\mu = 3000$ جالون وانحراف معياري $\sigma = 200$ جالون. أوجد الاحتمال أن المبيعات في الأسبوع ما بين 3200 و 3500 جالون .

الحل . إذا كان X متغير عشوائي يرمز لمبيعات الغاز الطبيعي في الأسبوع ، فان X يكون متغير عشوائي طبيعي متوسطه $\mu = 3000$ وانحرافه المعياري $\sigma = 200$.

عندما $x_1 = 3200$ فـإن :

$$z_1 = \frac{3200 - 3000}{200} = 1.0.$$

وعندما $x_2 = 3500$ فإن :

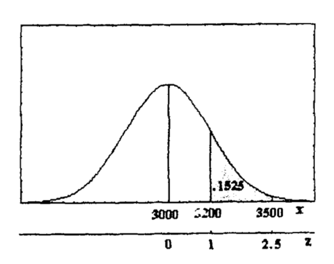
$$z_2 \frac{3500 - 3000}{200} = 2.5.$$

الاحتمال المطلوب هو :

$$P(3200 < X < 3500)$$

وهو يساوي المساحة المظللة في شكل (٦-١٣) أى أن :

$$P(3200 < X < 3500) = P(1.0 < Z < 2.5)$$
$$= P(0 < Z < 2.5) - P(0 < Z < 1)$$
$$= 0.4938 - 0.3413 = 0.1525.$$



شكل (٦-١٣)

مثال (٦-٢١) إذا كانت كمية المطر الذي يسقط سنويا في منطقة معينة متغير عشوائي يخضع للتوزيع الطبيعي بانحراف معياري $\sigma = 2$ بوصة. أوجد المتوسط السنوي لسقوط المطر في عام محدد إذا كان احتمال سقوط أكثر من 30 بوصة من المطر في هذا العام يساوي 0.0548 .

الحل . بفرض أن X متغير عشوائي طبيعي له متوسط غير معروف وانحراف معياري $\sigma = 2$، المتغير الطبيعي القياسي المناظر يكون :

$$Z = \frac{X - \mu}{\sigma}.$$

عندما $x_1 = 30$ فإن :

$$z_1 = \frac{30 - \mu}{2}.$$

ولكن احتمال أن X أكثر من 30 هو 0.0548، كما هو معطى، وعلى ذلك فان :

$$P(Z \geq \frac{30 - \mu}{2}) = 0.0548.$$

أي أن :

$$P(0 \leq Z \leq \frac{30 - \mu}{2}) = 0.4452.$$

ولكن من جدول التوزيع الطبيعي القياسي المعطى في ملحق (٣) فان :

$$P(0 \le Z \le 1.6) = 0.4452.$$

أي أن $\frac{30-\mu}{2} = 1.6$ ومنه $\mu = 26.8$ بوصة .

مثال (٦-٢٢) إذا كانت حموضة الدم الآدمي مقاس بدلالة الأس الأيدروجين متغير عشوائي طبيعي متوسطه $\mu = 7.2$. إذا كان احتمال أن يكون مستوى الأس الايدروجيني اكبر من 7.5 يساوي **0.0222** أوجد الانحراف المعياري للتوزيع .

الحل • عندما $x_1 = 7.5$ فإن :

$$z_1 = \frac{7.5-7.2}{\sigma} = \frac{0.3}{\sigma}.$$

أي أن :

$$P(Z \ge \frac{0.3}{\sigma}) = 0.0222.$$

وبالتالي فان :

$$P(0 \le Z \le \frac{0.3}{\sigma}) = 0.5 - 0.0222 = 0.4778.$$

ولكن من الجدول الطبيعي القياسي في ملحق (٣) نجد أن :

$$P(0 \le Z \le 2.01) = 0.4778.$$

أي أن $\frac{0.3}{\sigma} = 2.01$ ومنه $\sigma = 0.149$.

(٦-٦-٢) التوزيع الطبيعي كتقريب لذي الحدين

يتم الحصول على الاحتمالات المرتبطة بتجارب ذي الحدين مباشرة من الصيغة $b(x;n,p)$ أو من الجدول في ملحق (١) وذلك عندما تكون n صغيرة كما سبق ذكره في البند (٦-٢) . إذا كانت n غير موجودة في الجدول فإنه يمكن حساب احتمالات ذي الحدين بالطرق التقريبية . تناولنا في البند (٦-٤) كيفية استخدام توزيع بواسون لتقريب احتمالات ذي الحدين عندما تكون n كبيرة و p تقترب جدا من الصفر.

يعطي التوزيع الطبيعي بمتوسط $\mu = np$ وتباين $\sigma^2 = npq$ تقريب ممتاز لتوزيع ذي الحدين عندما تكون n كبيرة و p تقترب من 0.5 • في الحقيقة يمكن استخدام التوزيع الطبيعي بمتوسط $\mu = np$ وتباين $\sigma^2 = npq$ كتقريب لتوزيع ذي الحدين حتى ولو كانت n ليست كبيرة جدا و p ليست قريبة من 0 أو 1. وعلى أي حال، فان هذا التقريب يكون جيدا إذا

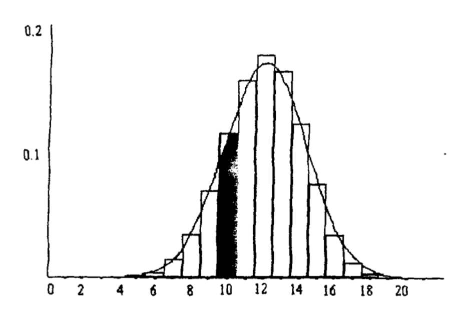
كانت كلا من np و nq أكبر من 5 . للتوضيح، بفرض أن X متغير عشوائي يتبع ذي الحدين حيث n=20 و p = 0.6. شكل (٦-١٤) يوضح المدرج الاحتمالي لهذا المتغير العشوائي وأيضا المنحني الطبيعي بمتوسط وانحراف معيارى:

$$\sigma=\sqrt{(20)(0.6)(0.4)}=2.19 \ , \ \mu=(20)(0.6)=12.$$

لمتغير عشوائي X يتبع توزيع ذي الحدين فإن P(X=x) يمثل مساحة المستطيل والذي تقع القيمة x في منتصف قاعدته وعلى ذلك يكون $b(x;n,p)$ مساوٍ تقريبا للمساحة تحت المنحنى الطبيعي الواقعة بين القيمتين $x-\frac{1}{2}$ و $x+\frac{1}{2}$. على سبيل المثال لحساب P(X=10) وباستخدام صيغة توزيع ذي الحدين فإن :

$$P(X=10)=\sum_{x=0}^{10} b(x;20,0.6)-\sum_{x=0}^{9} b(x;20,0.6)$$
$$=0.245-0128=0.117.$$

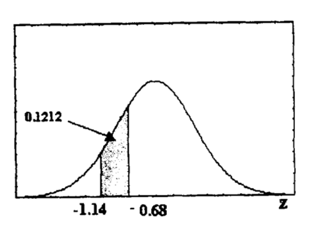
نفس الاحتمال السابق تقريبا يساوى المساحة المظللة تحت المنحنى الطبيعي بين $x_1=9.5$ و $x_2=10.5$ كما في شكل (٦-١٤).



شكل (٦-١٤)

أي أنه عندما $x_1=9.5$ فـــإن :

$$z_1=\frac{9.5-12}{2.19}=-1.14.$$



شكل (٦-١٥)

وعندما $x_2 = 10.5$ فإن :

$$z_2 = \frac{10.5-12}{2.19} = -.68.$$

إذا كان X متغير عشوائي يتبع ذي الحدين وZ متغير عشوائي يتبع التوزيع الطبيعي القياسي فإن :

$$P(X=10) = b(10;,20,0.6) \approx P(-1.14 \le Z \le -0.68)$$
$$= 0.3729 - 0.2517 \quad = 0.1212.$$

والتي تتفق مع القيمة المضبوطة التي حصلنا عليها باستخدام صيغة ذي الحدين (0.117)٠

مثال (٦-٢٣) إذا ألقيت زهرتا نرد 120 مرة وإذا كان اهتمامنا بمجموع النقط التي تظهر على السطح العلوي للنردين ٠ فما هو احتمال ظهور المجموع 7 على الأقل 15مرة٠

الحل ٠ بفرض أن المحاولات مستقلة وبما أن احتمال ظهور مجموع 7 في المحاولة الواحدة هو $\frac{6}{36} = \frac{1}{6}$ ٠ فإذا كان X متغير عشوائي يتبع ذي الحدين ويمثل عدد مرات ظهور المجموع 7 ٠ فإن x= 0,1,...,120 وعلى ذلك فان الاحتمال المطلوب هو :

$$P(X \ge 15) = \sum_{x=15}^{120} b(x;120,\frac{1}{6}).$$

و لصعوبة حساب قيمة هذا الاحتمال فإننا نلجأ إلى التوزيع الطبيعي ،كتقريب لتوزيع ذي الحدين، والذي متوسطه وانحرافه المعياري على التوالي :

$$\mu = np = (120)(\frac{1}{6}) = 20\,,$$

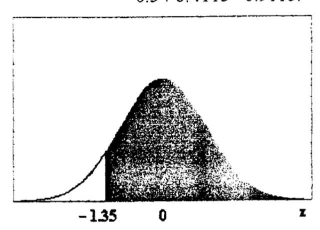
$$\sigma = \sqrt{npq} = \sqrt{(120)(\frac{1}{6})(\frac{5}{6})} = \sqrt{\frac{50}{3}} = 4.08.$$

للحصول على الاحتمال المطلوب فإننا نحتاج إلى حساب المساحة على يمين القيمة $x_1 = 14.5$ ، عندما $x_1 = 14.5$ فإن:

$$z_1 = \frac{14.5 - 20}{4.08} = -1.35$$

الاحتمال المطلوب يساوى المساحة المظللة في شكل (٦-١٦) ومن جدول التوزيـــع الطبيعـــي القياسي في ملحق (٣) فإن :

$$\begin{aligned} P(X \ge 15) &= \sum_{x=15}^{120} b(x;120,\frac{1}{6}) \approx P(Z \ge -1.35) \\ &= 0.5 + P(-1.35 \le Z \le 0) \\ &= 0.5 + P(0 \le Z \le 1.35) \\ &= 0.5 + 0.4115 = 0.9115. \end{aligned}$$



شكل (٦-١٦)

مثال (٦-٢٤) إذا كانت نسبة المصابين بارتفاع ضغط الدم في مجتمع كبير 4%. فـــإذا تم اختيار 500 شخصاً عشوائياً من هذا المجتمع وتم قياس ضغط دمهم، فما هو احتمال وجـــود 15 شخصا على الأقل مصابين بارتفاع ضغط الدم.

الحل . إذا كان X هو عدد الأشخاص المصابين بارتفاع ضغط الدم فانه يكون عبارة عن متغـــير عشوائي متقطع يتبع توزيع ذي الحدين حيث :

$$b(x;500,0.04)=\binom{500}{x}(0.04)^x\ (0.96)^{500-x}\qquad ,x=0,1,...,500.$$

والاحتمـــال المطلـــوب هـــو :

$$b(x\geq 15)=b(15;500,0.04)+b(16;500,0.04)+...+b(500;500,0.04)$$

ومن الصعب حساب قيمة هذا الاحتمال وبالتالي فإننا نلجأ إلى التوزيع الطبيعي، كتقريب لتوزيع ذي الحدين، والذي متوسطه وتباينه على التوالي هما :

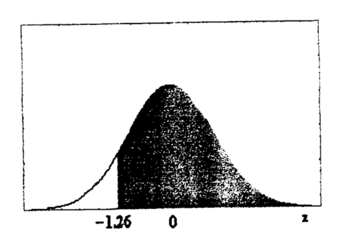
$$\mu = np=(500)(0.04)=20,$$

$$\sigma^2 = npq=(500)(0.04)(0.96)=19.2.$$

أي أن $\sigma = 4.38$ ، للحصول على الاحتمال المطلوب فإننا نحتاج إلى حساب المساحة على يمين القيمة $x_1 = 14.5$ ، عندما $x_1 = 14.5$ فإن :

$$z_1 = \frac{14.5-20}{4.38} = -1.26.$$

الاحتمال المطلوب يساوى المساحة المظللة في شكل (٦-١٧) ومن جدول التوزيع الطبيعـــي القياسي في ملحق (٣) فإن :



شكل (٦-١٧)

$$P(X\geq 15)\approx P(Z\geq -1.26)$$
$$=0.5+P(-1.26\leq Z\leq 0)$$
$$=0.5+P(0\leq Z\leq 1.26)$$
$$=0.5+0.3962=0.8962.$$

مثال (٦-٢٥) إذا كان من المعروف أن 6% من الأفراد الذكور مصابون بعمى الألوان . فإذا تم اختيار عينة من 200 فرد من الذكور وتم اختبارهم لمعرفة إصابتهم بعمى الألوان من عدمه . أوجد احتمال أن يكون عدد المصابين بعمى الألوان :

(أ) على الأقل 20 فردا (ب) على الأكثر 15 فردا (جـ) بالضبط 15 فردا

(د) $P(15 \leq X \leq 20)$.

الحل . (أ) سوف نستخدم التوزيع الطبيعي، كتقريب لذي الحدين، بمتوسط

$$\mu = np=(200)(0.06)=12$$

وتباين

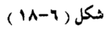
$$\sigma^2 =npq=(200)(0.06)(0.94)=11.28.$$

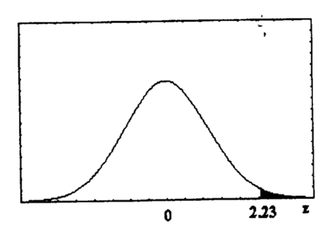
أي أن $\sigma = 3.36$. (ا) للحصول على الاحتمال $P(X\geq 20)$ فإننا نحتاج إلى حساب المساحة على يمين القيمة $x_1 = 19.5$. عندما $x_1 = 19.5$ فإن :

$$z_1 = \frac{19.5-12}{3.36} = 2.23.$$

الاحتمال المطلوب يساوى المساحة المظللة في شكل (٦-١٨) ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :

$$P(X\geq 20) = \sum_{x=20}^{200} b(x;200,0.06)$$
$$\approx P(Z\geq 2.23)$$
$$=0.5-P(0\leq Z\leq 2.23)$$
$$=0.5-0.4871=0.0129.$$


شكل (٦-١٨)

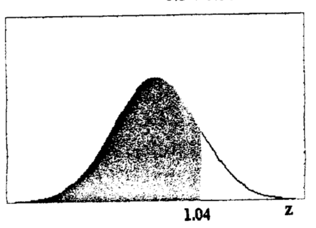


(ب) للحصول على الاحتمال $P(X \leq 15)$ فإننا نحتاج إلى حساب المساحة على يسار القيمة $x_1 = 15.5$ ، عندما $x_1 = 15.5$ فإن :

$$z_1 = \frac{15.5 - 12}{3.36} = 1.04.$$

الاحتمال المطلوب يساوى المساحة المظللة في شكل (٦-١٩) . ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :

$$\begin{aligned} P(X \leq 15) &= \sum_{x=0}^{15} b(x;200,0.06) \\ &\approx P(Z \leq 1.04) \\ &= 0.5 + p(0 \leq Z \leq 1.04) \\ &= 0.5 + 0.3508 = 0.8508. \end{aligned}$$



شكل (٦-١٩)

$$P(X=15)=\sum_{x=0}^{15} b(x;200,0.06)-\sum_{x=0}^{14} b(x;200,0.06).$$ (جـ)

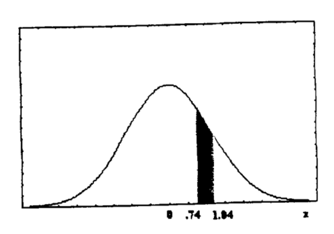
نفس الاحتمال السابق تقريبا يساوي المساحة تحت المنحنى الطبيعي بين $x_1 = 14.5$ و $x_2 = 15.5$ ، عندما $x_1 = 14.5$ فإن :

$$z_1=\frac{14.5-12}{3.36}=0.74.$$

وعندما $x_2 = 15.5$ فإن :

$$z_2=\frac{15.5-12}{3.36}=1.04.$$

الاحتمال المطلوب يساوي المساحة المظللة في شكل (٦-٢٠) ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :



شكل (٦-٢٠)

$$\begin{aligned}P(X=15)&\approx P(0.74\le Z\le 1.04)\\&=P(0\le Z\le 1.04)-P(0\le Z\le 0.74)\\&=0.3508-0.2704=0.0804.\end{aligned}$$

$$P(15\le X\le 20)=\sum_{x=0}^{20} b(x;200,0.06)-\sum_{x=0}^{14} b(x;200,0.06).$$ (د)

نفس الاحتمال السابق تقريبا يساوي المساحة تحت المنحنى الطبيعي بين $x_1 = 14.5$ و $x_2 = 20.5$ ،

عندما $x_1 = 14.5$ فإن :

$$z_1=\frac{14.5-12}{3.36}=0.74.$$

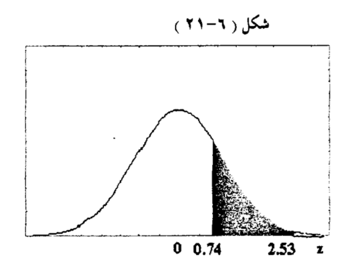
وعندما $x_2 = 20.5$ فإن :

$$z_2 = \frac{20.5 - 12}{3.36} = 2.53.$$

الاحتمال المطلوب يساوى المساحة المظللة في شكل (٦-٢١) ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :

$$P(15 \le X \le 20) \approx P(0.74 \le Z \le 2.53)$$
$$= P(0 \le Z \le 2.53) - P(0 \le Z \le 0.74)$$
$$= 0.4943 - 0.2704 = 0.2399.$$

شكل (٦-٢١)



تمارين

-١- إذا ألقيت زهرة نرد متزنة مرة واحدة، وإذا كان X يمثل عدد النقط التي تظهر على الوجه العلوي للزهرة عند الرمى. المطلوب (ا) التوزيع الاحتمالي للمتغير العشوائي X.

(ب) $P(X > 4)$, $P(X - 3 < 0)$ (جـ) التوقع والتباين للمتغير العشوائي X .

(د) التمثيل البياني للتوزيع الاحتمالي للمتغير العشوائي X .

-٢- فصل دراسي يتكون من 70 تلميذا بحيث يمثل كل تلميذ عددا من واحد إلى 70 . بفرض أنه تقرر اختيار تلميذ عشوائيا، وإذا كان X يمثل العدد الذي يتم إختياره. أوجد الاحتمالات الآتية (ا) $P(X \le 8)$ (ب) $P(5 \le X \le 30)$.

-٣- أوجد صيغة للتوزيع الاحتمالي الخاص بالمتغير العشوائي X الذي يمثل رقم الكـــرة المختـــارة عشوائيا من وعاء يحتوى على 10 كرات مرقمة من واحد إلى 10 ، ما هو الاحتمـــال أن يكـــون الرقم المختار أقل من 5 ؟

-٤- أوجد التوزيع المنتظم للعينات من الحجم n=3 المراد اختيارها من مجتمع حجمه N=6 ،

-٥- أوجد التوزيع الاحتمالي للفئات الجزئية من الأشهر من الحجم n=3 ،

-٦- إذا كان X متغير عشوائي منفصل يتبع التوزيع المنتظم، أي أن

$$f(x;N)=\frac{1}{N},x=1,2,\ldots,N$$

أوجد E(X) وتباين X ،

-٧- إذا اختير رقما عشوائيا من بين الأعداد 0,1,...,9 (ا) أوجد التوزيع الاحتمالي للعـــدد الذي يظهر (ب) التمثيل البياني لهذا التوزيع ،

-٨- لوحظ لفترة طويلة أن صياد يصيب الهدف باحتمال 0.9 ، فإذا أطلق الصياد 4 طلقات على هدف، أوجد احتمال (ا) إصابة الهدف 3 مرات؟ (ب) إصابة الهدف مرتين على الأكثر ،

-٩- خلال فترة طويلة من الزمن وجد أن فاعلية دواء في شفاء الحالات التي يوصف لها هو 20% ، إذا وصف هذا الدواء لأربعة مرضى ، أوجد احتمال (أ) يكون له تأثير في الشفاء لثلاثة مرضى على الأقل ،

-١٠- أثبتت إحصائيات إحدى شركات التأمين أن 0.002 من الحوادث المســـجلة في الشـــركة خطيرة ، احسب : (ا) احتمال عدم وجود حوادث خطيرة في 40 حادثة مســـجلة تم اختيارهـــا عشوائيا (ب) احتمال وقوع ثلاثة حوادث خطيرة في 20 حادثة مسجلة تم اختيارها عشوائيا ،

-١١- احتمال أن تصيب أي طائرة أحد أهداف العدو هو 0.9، فإذا أغارت خمس طائرات علـــى الهدف أوجد (ا) التوزيع الاحتمالي لعدد الطائرات التي تصيب الهدف (ب) متوســـط التوزيـــع وكذلك الانحراف المعيارى ،

-١٢- إذا كان 90% من الطلاب في جامعة ما ينجحون في مادة الإحصاء، فما احتمـــال فشـــل أثنين في فصل يتضمن ثلاثين طالبا ؟

-١٣- إذا كانت المحركات الأربعة لطائرة تعمل مستقلة بعضها عن بعض ، فإن احتمال تعطـــل أي منها والطائرة في الجو هو 0.5 ، ما احتمال أن يكون الطيران ناجح إذا كانت عملية الطيران تحتاج لمحرك واحد على الأقل ؟

-١٤- إذا كان احتمال أن 4% من الوحدات في مصنع ما تالفة ، ما هو احتمال إنه على الأكـــثر توجد وحدة تالفة في عينة عشوائية من الحجم n=30 ؟

-١٥- احتمال أن يكسب فريق ما في أي مباراة يلعبها 0.75 فإذا كان الفريق سوف يلعب 8 مباريات أوجد احتمال أن يكسب (ا) مباريتين بالضبط (ب) على الأقل مباراة واحدة (ج) أكثر من نصف المباريات.

-١٦- أسرة بها خمسة أطفال. أوجد احتمال أن يكون بينهم (ا) 3 أولاد (ب) عدد الأولاد أقل من عدد البنات وذلك تحت فرض أن احتمال أن يكون الطفل ولد هو 0.5 .

-١٧- احتمال أن لاعب كرة القاعدة يحصل على ضربة في أي وقت يلعب فيه هو 0.3 . فإذا كان سوف يلعب 100 مباراة في الشهر القادم أوجد المتوسط والتباين للمتغير العشوائي X .

-١٨- إذا كانت فاعلية مبيد معلن عنه في قتل الحشرات في رشة واحدة هو 90% فإذا كان هناك 10,000 حشرة سوف تعامل برشة واحدة من هذا المبيد. أوجد التوقع الرياضي والتباين والانحراف المعياري لعدد الحشرات التي تقتل.

-١٩- قدرت شركة للطيران احتمال وصول طائرتها في ميعادها والتي تقوم من البلد A إلى البلد B هو 0.96 . فإذا قامت خمس طائرات لهذه الشركة من مطار البلد A متجه إلى البلد B أوجد (ا) التوزيع الاحتمالي لعدد الطائرات التي تصل في ميعادها (ب) احتمال وصول طائرتين في ميعادهم (جـ) التمثيل البياني لهذا التوزيع.

-٢٠- اشترى تاجر عشر ثلاجات. فإذا كان احتمال أن تكون أي منها تالفة هو 0.1 فما هو احتمال إن يكون من بينها (ا) ثلاجتين تالفتين (ب) ثلاجتين على الأقل تالفتين (جـ) أربعة على الأقل تالفة.

-٢١- تعمل 10 ماكينات في مصنع، فإذا كان احتمال أنه في نهاية اليوم سوف تعطل الماكينة هو 0.2 . وإذا كانت الماكينات تعمل بصورة مستقلة عن بعضها أوجد التوزيع الاحتمالي لعدد الماكينات التي تعمل حتى نهاية اليوم مع تمثيل التوزيع الاحتمالي بالمدرج التكراري.

-٢٢- بفرض أنه يوجد من بين كل 400 سيارة من إنتاج مصنع معين لإنتاج السيارات 20 سيارة غير سليمة، أي غير صالحة للاستعمال. سحبت عينة عشوائية مكونة من 3 سيارات من إنتاج ذلك المصنع، أوجد احتمال أن يكون من بينها سيارتين غير سليمة.

-٢٣- إذا كان احتمال حصول أبوين على طفل أشقر الشعر هو 0.25، فإذا كان في الأسرة 8 أطفال. ما احتمال أن يكون نصفهم ذو شعر أشقر.

-٢٤- يأخذ مراقب جودة إنتاج عينة عشوائية من 10 صمامات عشوائيا من شحنة كبيرة من الصمامات معروف عنها أنها تحتوى على 30% صماما معيبا. ما هو احتمال أن يكون عدد الصمامات المعيبة في العينة أكثر من أو يساوى 2 ؟

-٢٥- إذا كان احتمال إصابة هدف بقذيفة واحدة هو 0.2، ما هو احتمال إصابة الهدف مرتين على الأقل خلال 3 قذائف ؟

-٢٦- احتمال كشف جهاز رادار لطائرة معادية هو 0.99، فإذا كان لدينا 4 أجهزة تعمل مستقلة بعضها عن بعضها، المطلوب إيجاد (ا) ظهور طائرة معادية في سمائنا (ب) ظهور طائرة معادية على شاشات أربعة منها .

-٢٧- إذا كان 5% من نوع معين من صمامات التلفزيون يحترق قبل انتهاء مدة الكفالة . فإذا بيع ألف صمام فما هو متوسط وتباين X ، حيث X تمثل عدد الصمامات المحترقة قبل مدة كفالتها ؟

-٢٨- إذا كان احتمال تحمل نوع معين من المصابيح للضغط العالي هو 0.1 فإذا أخذنا عينة من 200 مصباح فما هو احتمال ألا يتحمل 20 منها للضغط العالي؟

-٢٩- في مدينة كبيرة كان عدم الوفاق بين الزوجين سببا لـ 60% من حالات الطلاق . أوجد احتمال أن أربعة من حالات الطلاق الستة القادمة في المدينة سيعزى إلى هذا السبب .

-٣٠- في بحث ميداني في بلد ما وجد أن 2% من الأشخاص يفضلون شراء تلفزيون بلون أبيض . ما هو احتمال أن أكثر من نصف التلفزيونات التي سوف تباع من بين 40 تلفزيون لونها أبيض ؟

-٣١- يحتوى امتحان على 18 سؤال من نوع صح وخطأ، كل سؤال له أربع أجوبة محتملة منهم فقط واحد هو الإجابة الصحيح . ما هو احتمال أن شخص يمكنه الإجابة على 5 إجابات صحيحة وذلك بطريقة التخمين ؟

-٣٢- احتمال أن يشفى مريض من عملية في القلب هو 0.9 ، ما هو احتمال أن يشفى 5 مرضى من بين 10 مرضى سوف يجرى لهم العملية ؟

-٣٣- يحتوى امتحان على 20 سؤال من نوع صح وخطأ، كل سؤال له أربع أجوبة محتملة منهم فقط واحد هو الإجابة الصحيحة . فإذا كان احتمال أن يحصل طالب على 100% في الاختبار هو $P(100\%) = (\frac{1}{4})^{20}$. المطلوب إيجاد احتمال حصوله على 80% إجابة صحيحة .

-٣٤- إذا كان 0.77 من المستهلكين يستخدمون مسحوق ما . ما هو الاحتمال في عينة من 10 مستهلكين (ا) بالضبط أربعة يستخدمون المسحوق (ب) على الأقل أربعة يستخدمون المسحوق (جـ) ليس أكثر من واحد يستخدم المسحوق (د) 8 أو أكثر يستخدمون المسحوق .

-٣٥- أثبتت الإحصائيات في إحدى المستشفيات أن 40% من المرضى الذين يتلقون العلاج بها يفشلون في دفع فاتورة الحساب . بفرض أن 4 مرضى دخلوا المستشفى لتلقى العلاج، أوجد

الاحتمالات التالية (ا) كل المرضى لن يدفعون الفاتورة (ب) واحد على الأقل سوف يدفع الفاتورة .

-٣٦- إذا كان 90% من المتقدمين لوظيفة دبلوماسي يفشلون في المقابلة الأولى، فإذا اختير 12 شخصا عشواليا من المتقدمين، أوجد احتمال انه (ا) على الأقل 5 يفشلون (ب) من10 إلى 12 يفشلون .

-٣٧- إذا كان X متغيرا عشوائيا يمثل عدد الطلبة الذين يحصلون على وظائف بعد سنة واحدة من التخرج يتبع توزيع ذي الحدين بمعلمتين n=250 , p=0.98 (ا) أوجد التوزيع الاحتمالي للمتغير العشوائي X (ب) أوجد المتوسط والتباين للتوزيع الاحتمالي للمتغير X .

-٣٨- أثبتت التجارب حدوث تلف في التمثيل الغذائي لطفل واحد من بين 100 يولدون . بفرض أنه تم ولادة أربعة أطفال في يوم معين، ما هو احتمال (ا) ليس اكثر من طفل واحد لديه تلف (ب) عدم حدوث تلف .

-٣٩- بفرض أن 20% من سائقي سيارات الأجرة الذين يصلون إلى موقف السيارات يعطون إشارات حمراء في كل الاتجاهات (يتبعون الأنظمة) فإذا اختير 20 سائقا عشوائيا من الذين تم وصولهم إلى موقف السيارات . أوجد احتمال (ا) على الأكثر6 يتبعوا النظام (ب) بالضبط 10يتبعون النظام (جـ) على الأقل 12 يتبعون النظام .

-٤٠- استخدم جدول ذي الحدين في جدول ذي الحدين في ملحق (١) في إيجاد الاحتمالات

$P(x=4), n=15, p=0.4$

$P(x \leq 4), n=20, p=0.1$

$P(5 \leq X \leq 11), n=20, p=0.2$

$P(x=6), n=15, p=0.5$

$P(x \geq 8), n=15, p=0.5$

-٤١- في إحدى الدراسات لنزلاء مؤسسات الأحداث وجد أن 90% من النزلاء قد يعودون مرة ثانية إلى المؤسسة بعد انتهاء مدقمم . فإذا أخذت عينة عشوائية مكونة 5 من نزلاء من إحدى المؤسسات ، أوجد (ا) التوزيع الاحتمالي لعدد النزلاء الذين يعودون مرة أخرى إلى المؤسسة . (ب) عدد النزلاء المتوقع عودقم (ب) الانحراف المعياري لعدد النزلاء المتوقع عودقم مرة أخرى إلى المؤسسة .

-٤٢- إذا كان X متغيرا عشوائيا يمثل عدد المرضى الذين يلغون حجزهم عند طبيب الأسنان، فإذا كان معالم توزيع ذي الحدين للمتغير X هو n=15 , p=0.05 ، أوجد احتمال (ا) ثلاثة أشخاص سوف يلغون الحجز (ب) على الأكثر أربعة سوف يلغون الحجز .

-٤٣- في تجربة زراعية كانت نسبة الإصابة بفطر ما 0.2 في نهاية التجربة، فإذا كان X متغيرا عشوائيا يمثل عدد النباتات المصابة بالفطر في عينة من 10 نباتات، أوجد (ا) التوزيع الاحتمالي لعدد النباتات المصابة (ب) أوجد $P(X=0)$.

-٤٤- يولد30% من مواليد سلالة معينة من الأرانب بشعر طويل أوجد (ا) التوزيع الاحتمالي لعدد الحيوانات التي تولد بشعر طويل في بطن من أربعة أرانب (ب) أوجد $P(X=1)$ (جـ) أوجد المتوسط والتباين للتوزيع الاحتمالي للمتغير X .

-٤٥- ينبت75% من شاتلات الأشجار المزروعة بشركة معينة لتشجير الطرق أوجد المتوسط والتباين للتوزيع الاحتمالي للمتغير X والذي يمثل عدد الأشجار التي تنبت من مجموعة مكونة من10 أشجار مزروعة.

-٤٦- أعطيت ستة فئران جرعة معينة من سم ولوحظ عدد الفئران التي تموت خلال 72ساعة، فإذا كان احتمال موت كل فأر هو 0.2 ،أوجد التوزيع الاحتمالي لعدد الفئران التي تموت خلال 72 ساعة وأوجد المتوسط والتباين للمتغير العشوائي X .

-٤٧- إذا كان نسبة الأميين في إحدى القرى هي 33%، فإذا أخذت عينة عشوائية مكونة من5 أشخاص من أفراد هذه القرية، أوجد التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل عدد الأميين في العينة وأوجد المتوسط والتباين للمتغير العشوائي X .

-٤٨- درب حيوان على لمس واحدة من رافعتين إذا أمر بذلك، بفرض أن احتمال أن يلمس الرافعة الصحيحة إذا أمر بذلك هو 0.8 أوجد التوزيع الاحتمالي لعدد مرات لمس الرافعة الصحيحة في محاولات عددها 10 وأوجد المتوسط والتباين للتوزيع الاحتمالي الذي حصلت عليه.

-٤٩- وجد أن فصيلة دم 40% من سكان مدينة ما تكون من نوع معين، ما احتمال أن يكون ثلاثة فقط من مجموعة مكونة من ثمانية أشخاص تم اختيارهم عشوائيا من سكان هذه المدينة لهم فصيلة الدم هذه؟

-٥٠- بالرجوع إلى التاريخ العائلي لزوجين وجد أن احتمال أن يكون أيا من أطفالهم مصابا بعيب خلقي معين 0.05 فإذا كان للزوجين أربعة أطفال أوجد التوزيع الاحتمالي لعدد الأطفال المصابين بتخلف عقلي وأوجد المتوسط والتباين للمتغير العشوائي X.

-٥١- إذا كان احتمال ولادة طفل أعسر فى بلد ما، هو 0.01 أوجد التوزيع الاحتمالي لعدد الأطفال الذين لديهم هذه الصفة وأوجد المتوسط والتباين للمتغير العشوائي X.

-٥٢- إذا كان احتمال التحاق إحدى الخريجين بالعمل فور تخرجه هو0.75 . فإذا اختيرت عينة عشوائية من 30 فرد من حديثي التخرج، أوجد (ا) احتمال عدم التحاق أي شخص بالعمل فور

تخرجه (ب) احتمال التحاق أثنين على الأقل بالعمل فور تخرجهم (جـ) العدد المتوقع للأشخاص الذين يلتحقون بالعمل فور تخرجهم.

-٥٣- في عملية تصنيع كرات التحميل وجد أن احتمال وجود كرة تالفة 0.1 ما هو احتمال الحصول على 10 كرات تحميل تالفة من عينة عشوائية بها 1000 وحدة؟

-٥٤- إذا كان طرد بريدي يمكن أن يفقد أو يتلف أو يصل إلى صاحبه. فإذا كان احتمال أن يتلف 0.2 واحتمال أن يفقد هو 0.4 واحتمال أن يصل إلى صاحبه هو 0.4 . فإذا أرسل 15 طردا إلى بلد ما هو احتمال أن يصل 13 منهم بأمان إلى أصحابهم ويفقد 1 ويتلف 1.

-٥٥- تبعا لنظرية الوراثة فإن نوع معين من الحيوانات تلد حيوانات لونها أحمر وأبيض و أسود بنسبة 4 : 4 : 8 أوجد احتمال أنه من بين 8 مواليد سوف يكون منهم 5 لونهم أحمر و 2 لونهم أسود وواحد لونه أبيض.

-٥٦- صندوق به 15 ثمرة فاكهة، منها 4 تالفة، أوجد التوزيع الاحتمالي لعدد الثمار التالفة ثم مثل بيانيا الدالة الإحتمالية و n=3

-٥٧- يوجد في مكتبة 20 نسخة من كتاب في مقدمة الإحصاء، منهم 12 طبعة أولى و 8 طبعة ثانية. فإذا اختيرت عينة عشوائية من الحجم n=5 ، وإذا كانت X تمثل عدد الكتب المختارة من الطبعة الثانية، أوجد $P(X=2)$.

-٥٨- صندوق يحتوى على رقائق بالغة الصغر منها 10 جيدة و3 تالفة. فإذا تقرر اختيار عينة عشوائية من ثلاثة رقائق . وإذا كانت X تمثل عدد الوحدات التالفة في العينة المختارة، أوجد $P(X \leq 3)$.

-٥٩- قام باحث في علم الجيولوجيا بتجميع 10 وحدات من صخور البازلت و 10 وحدات من الجرانيت، فإذا كان للباحث معمل وطلب من مساعده أن يختار عينة عشوائية من15 وحدة للتحليل أوجد (ا) التوزيع الاحتمالي لعدد وحدات البازلت في العينة المختارة (ب) احتمال أن الوحدات المختارة من نفس النوع.

-٦٠- إذا كان عدد الجزيئات المشعة من مصدر إشعاعي يتبع توزيع بواسون وإذا كان $P(X=0)=0.3$ فما احتمال إشعاع جزيئين أو اكثر ؟

-٦١- وعاء به جزئ 10000 فإذا كان احتمال هروب جزئ من الوعاء هو 0.0004 فما احتمال هروب أكثر من 5 جزيئات.

-٦٢- تمر في المتوسط 20 سيارة في الدقيقة من أمام كشك رسوم المرور خلال ساعة الذروة. أوجد احتمال مرور 7 سيارات من أمام الكشك خلال دقيقة مختارة عشوائيا.

-٦٣- إذا كانت نسبة المعيب في مصنع مصابيح كهربائية هو 0.1 ، فإذا اخترنا عينة عشوائية من 200 مصباح، أوجد احتمال أن يوجد في هذه العينة مصباحا واحدا على الأكثر معيب وذلك باستخدام توزيع بواسون كتقريب لذي الحدين.

-٦٤- إذا كان متوسط وقوع الزلازل في بلد ما هو3 في السنة ، أوجد احتمال أن يقع زلزالا واحدا على الأقل في هذه السنة.

-٦٥- إذا كانت متوسط عدد الحوادث الشهرية في إحدى الطرق هو 0.2 . اختير شهرا عشوائيا، أوجد احتمال وقوع حادثتين على الأقل.

-٦٦- تقوم إحدى المصانع بإنتاج منتج معين معبا في أكياس عبوة الكيس الواحد كيلوجرام. فإذا كان احتمال وجود كيس واحد غير مطابق للموصفات هو 0.05 ، فإذا اختير عينة عشوائية من 20 كيس وبافتراض أن X متغيرا عشوائيا يمثل عدد الأكياس الغير مطابقة في العينة. أوجد (ا) التوزيع الاحتمالي لعدد الأكياس الغير مطابقة للمواصفات في العينة (ب) احتمال وجود كيسين على الأقل غير مطابقين للمواصفات (جـ) احسب العدد المتوقع للأكياس الغير مطابقة للمواصفات.

-٦٧- في المتوسط يتوقف 7 عملاء للتزود بالوقود عند طلمبة بنزين في الساعة. ما هو احتمال (ا) توقف 7 عملاء في الساعة (ب) أربعة عملاء أو أقل في ساعة ما.

-٦٨- تشير الدراسات على أن 0.002 من القوى العاملة القومية في بلد ما يصابون بمرض خطير خلال عام. فإذا اختير شخص عشوائيا. أوجد (ا) القيمة المتوقعة لعدد الذين يمرضون في العام (ب) احتمال أن عاملين يمرضون خلال عام. استخدم توزيع بواسون كتقريب ذي الحدين.

-٦٩- بفرض أن متغيرا عشوائيا يمثل عدد الجرائم التي تحدث في بلد ما في الفترة ما بين الساعة الواحدة صباحا حتى الساعة الثانية يتبع توزيع بواسون حيث $\mu = 0.2$ (ا) أوجد التوزيع الاحتمالي للمتغير العشوائي X (ب) أوجد $P(X = 0)$.

-٧٠- إذا كان عدد التذاكر التي تصرف في موقف للسيارات في الساعة يتبع توزيع بواسون بمعلمة $\mu = 4$ (ا) ما هو الاحتمال أنه بالضبط تصرف 3 تذاكر خلال ساعة معينة (ب) ما هو الاحتمال أنه على الأقل تصرف 4 خلال ساعة معينة (جـ) ما هو عدد التذاكر المتوقع صرفها خلال 45 دقيقة.

-٧١- تصل طائرة شراعية إلى المطار بمتوسط $\mu = 8$ في الساعة، ما هو الاحتمال (ا) خمسة بالضبط سوف يصلون خلال ساعة (ب) على الأقل خمسة يصلون خلال ساعة (جـ) على الأقل 10 يصلون خلال ساعة (د) ما هي القيمة المتوقعة والانحراف المعياري لعدد الطائرات التي تصل إلى المطار خلال 90 دقيقة(هـ) ما هو الاحتمال أنه على الأقل تصل طائرة خلال 2.5 ساعة ؟

-٧٢- في بحث وجد أن شخص واحد من 1000 شخص يحمل جين تالف يـــؤدى إلى الإصابـــة بسرطان القولون . اختيرت عينة عشوائية من 15 شخص أوجـــد (ا) التوزيـــع التقريـــبي لعـــدد الأشخاص الذين يحملون الجين التالف (ب) استخدم هذا التوزيع فى حساب الاحتمال التقريبي أن 7 يحملون الجين التالف .

-٧٣- قامت شركة لبيع السيارات بإرسال طلب لكل مشتر منها سيارة معينة وذلك لفحصها من العيوب . بفرض أن 0.001 يمثل احتمال وجود عيب في السيارة . إذا اختيرت عينـــة عشـــوائية 10000 أوجد (ا) القيمة المتوقعة لعدد السيارات في العينة التي بها هذا العيب (ب) ما هو الاحتمـــل التقريبي أنه على الأقل 10سيارات بها هذا العيب (جـ) ما هو الاحتمال التقريـــبي لعـــدم وجـــود سيارات في العينة بها عيوب .

-٧٤- يستقبل سنترال في المتوسط ثلاث مكالمات في الدقيقة وذلك من أماكن بعيدة والمطلـــوب (ا) احتمال عدم وصول أي مكالمات في دقيقة واحدة (ب) احتمال وصول مكـــالمتين في دقيقتـــين (جـ) احتمال وصول أربع مكالمات في ثلاث دقائق .

-٧٥- إذا كان X متغير عشوائي يمثل عدد الأشخاص الذين يصلون إلى نادى رياضي بمتوسـط $\mu = 2$ في الدقيقة المطلوب (ا) احتمال وصول شخص واحد في دقيقـــة واحـــدة (ب) احتمـــال وصول 5 أشخاص في ثلاثة دقائق (جـ) احتمال عدم وصول أي شخص في دقيقة واحدة .

-٧٦- يستقبل سنترال في المتوسط ثلاثة مكالمات في الدقيقة وذلك من أماكن بعيدة والمطلـــوب (ا) احتمال عدم وصول أي مكالمات في دقيقة واحدة (ب) احتمال وصول مكـــالمتين في دقيقتـــين (جـ) احتمال وصول 4 مكالمات في ثلاث دقائق .

-٧٧- تقوم ماكينة بتصنيع الأقمشة، فإذا كان في المتوسط يوجد عيب لكل 10 ياردة من القماش أوجد (ا) احتمال عدم وجود عيوب في 2 ياردة من القماش (ب) احتمال وجود عيبين في ياردة مـن القماش .

-٧٨- إذا كان عدد المكالمات التي يتلقاها مركز للشرطة بين 8:30 و8:00 مساء فى يوم الجمعـــة يتبع توزيع بواسون بمعلمة $\mu = 2$ أوجد (ا) احتمال عدم وصول مكالمتين خلال هذه الفـــترة (ب) احتمال وصول مكالمتين خلال هذه الفترة (جـ) احتمال وصول ثلاث مكالمات أو أقل خلال هـــذه الفترة .

-٧٩- قررت عائلة تنظيم الإنجاب إذا رزقها الله بخمسة ذكور . فإذا كان احتمال ولادة ذكر في هذه العائلة هو 0.4 . أوجد التوزيع الاحتمالي لعدد مرات الحمل (الوضع) .

-٨٠- قررت أسرة أن تنظم النسل إذا رزقها الله بثلاثة أطفال من نفس الجنس. بفرض أن احتمال ولادة طفل ذكر تساوى احتمال ولادة طفل أنثى تساوى 0.5 . أوجد التوزيع الاحتمالي لعدد مرات الحمل.

-٨١- يلعب فريق A مع الفريق B سلسلة من المباريات. فإذا كان احتمال أن A يكسب في المباراة الواحدة التي يلعبها 0.6 وإذا كانت المباريات مستقلة، أوجد احتمال أن الفريق A قد يكون كسب أربعة مباريات في المحاولة السادسة.

-٨٢- أوجد احتمال أن شخص يلقى عملة سوف يحصل على صورة في المحاولة السابعة.

-٨٣- احتمال أن طالب يجتاز امتحان للحصول على رخصة قيادة طائرة هو 0.7 أوجد احتمال أن الشخص ينجح في الامتحان (أ) في المحاولة الثالثة (ب) قبل المحاولة الخامسة.

-٨٤- إذا كانت الدرجات إلى حصل عليها 100 طالب تتبع التوزيع الطبيعي بمتوسط 60 درجة وتباين 0.5. اختير طالبا بطريقة عشوائية (ا) ما هو احتمال أن تزيد درجته عن 72 ؟ (ب) ما هو عدد الطلاب الذين تقل درجاتهم عن 56 ؟

-٨٥- إذا حصل 10% من الطلاب على جوائز بسبب ارتفاع درجاتهم فما هي أدنى درجة يجب أن يحصل عليها الطالب حتى يحصل على جائزة ؟

-٨٦- إذا كان التوزيع التكراري لضغط الدم طبيعيا وكان متوسط الضغط الطبيعي هو 120 سم من الزئبق والانحراف المعياري هو 15 سم من الزئبق .

(أ) فما هو نسبة الأشخاص المحتمل أن يكون ضغطهم 150 فأكثر ؟

(ب) إذا علمت أن احتمال الحصول على شخص ضغط دمه أقل من قيمة معينة هي x_1 مثلا هو 0.8942 ، فما هي قيمة x_1؟

-٨٧- إذا كان الزمن اللازم لهضم وحدة واحدة من طعام معين يتبع التوزيع الطبيعي بمتوسط قدره 25 دقيقة وانحراف معياري قدره 3 دقائق . ما هو احتمال أن تهضم وحدة طعام في أقل من 30 دقيقة ؟

-٨٨- أوزان الدجاج في مزرعة ما تتبع التوزيع الطبيعي بمتوسط 6 رطل وانحراف معياري 1.5 أوجد (أ) احتمال أن يكون وزن الدجاجة اختيرت عشوائيا اكبر من سبعة أرطال .

(ب) نسبة الدجاج الذي وزنه اقل من أربعة أرطال .

-٨٩- تنتج ماكينة للمشروبات الباردة في المتوسط7 أوقيات من العصير لكل كوب. بفرض أن كمية الشراب يتبع التوزيع الطبيعي بانحراف معياري 0.5 أوقية . المطلوب (أ) نسبة الأكواب التي تحتوي على الأقل 7.8 أوقية . (ب) ما هو احتمال أن كوب يحتوي من بين6.7 إلى 7.9 أوقية .

-٩٠- يقطع شخص المسافة من مزله إلى عمله يوميا في زمن قدره 24 دقيقة في المتوسط بانحراف معياري 3.8 دقيقة . بفرض أن الزمن الذي يستغرقه يوميا يتبع التوزيع الطبيعي . أوجد احتمال أن الزمن الذي يستغرقه على الأقل نصف ساعة .

-٩١- إذا كانت أطوال 1000 طالبا تتبع التوزيع الطبيعي بمتوسط 68.5 بوصة وانحراف معياري 2.7 بوصة. كم عدد الطلبة المتوقع أن تكون أطوالهم : (أ) اقل من 63 بوصة . (ب) بين 67.5 الى 71 بوصة . (جـ) تساوي 69 بوصة . (د) اكبر من 74 بوصة .

-٩٢- تدفع شركة أجور العاملين فيها بمتوسط 100 جنيه لكل ساعة بانحراف معياري 5 جنيه. إذا كانت الأجور تقريبا تتبع التوزيع الطبيعي. ما هى النسبة المئوية من للعاملين الذين أجورهم تتراوح بين 80 إلى 90 في الساعة ؟

-٩٣- إذا كان معدل الذكاء I.Q لمجموعة من الطلبة الراغبين في الالتحاق في جامعة ما يتبع التوزيع الطبيعي بمتوسط $\mu=115$ وانحراف معياري $\sigma=12$.أوجد احتمال أن يكون معدل الذكاء أكبر من 120.

-٩٤- إذا كان متوسط العمر لمولد كهربائي صغير هو 10 سنوات بانحراف معياري 2 سنة. أوجد احتمال أن يقل عمر المولد عن 8 سنوات .

-٩٥- إذا كانت درجة الحرارة في مدينة ما يتبع التوزيع الطبيعي بمتوسط 20 درجة وانحراف معياري 3 درجات. أوجد احتمال أن تكون درجة الحرارة في أحد الأيام اقل من 25 درجة .

-٩٦- إذا كان قطر السلك الكهربى من إنتاج شركة ما يتبع التوزيع الطبيعي بمتوسط 8 ملليمتر وتباين 0.0004 ملليمتر. اشتري شخص سلك فما هو احتمال أن قطره لا يتعدي 8.0 ملليمتر .

-٩٧- إذا كان متوسط أطوال مجموعة من الجنود في معركة هو 70 بوصة. وإذا كان 10% من الجنود أطول من 72بوصة. بفرض أن أطوال الجنود في هذه المعركة يتبع التوزيع الطبيعي ما هو الانحراف المعياري ؟

-٩٨- إذا كان أطوال مجموعة من الجنود يتبع التوزيع الطبيعي. إذا كان % 13.75 منهم أطول من 72.2بوصة و % 8.08 أقصر من 67.2 بوصة. ما هو المتوسط والانحراف المعياري لأطوال الجنود ؟

-٩٩- إذا كان أوزان الذكور البالغين يتبع التوزيع الطبيعي وأن % 6.6 تحت 130 رطل وأن % 77.45 بين 130 إلى 180 رطل. أوجد معالم التوزيع .

-١٠٠- إذا كان دخول الأسر في مدينة ما يتبع التوزيع الطبيعي بمتوسط $\mu=15000\$$ والانحراف معياري $\sigma=3000\$$. ما هو احتمال أن يكون دخل أسرة اختيرت عشوائيا بين :

(أ) من 16000$ إلى 18000$ (ب) اقل من 12000$ (جـ) اعلي من 15000$.

-١٠١- تتبع درجات اختبارات الذكاء لمطوعي الجيش في سنة ما التوزيـــع الطبيعـــي بمتوســط $\mu=110$ وانحراف معياري $\sigma=10$. ويريد الجيش أن يعطي تدريبا متقدما لأعلى 10% من درجـات اختبارات الذكاء. ما هي أقل درجة في اختبارات الذكاء التي تقبل لحضور التدريب المتقدم ؟

-١٠٢- إذا ألقيت زهرة نرد 400 مرة. استخدم التوزيع الطبيعي كتقريب لذي الحدين لإيجــاد احتمال ظهور الرقم واحد (أ) 185 إلى 210 مرة (ب) بالضبط 205 مرة.

(جـ) أقل من 176 مرة .

-١٠٣- في مدينة ما وجد أن 10% من المدخنين مصابون بالسرطان. أخذت عينة من هذه المدينـة من 300 مدخن وفحصوا للتحقق من إصابتهم بالسرطان المطلوب:

(أ) احتمال أن تحوى العينة على 25 شخص مصاب بالسرطان .

(ب) ستون شخصا على الأقل مصابا بالسرطان .

(باستخدام التوزيع الطبيعي كتقريب لذي الحدين) .

-١٠٤- ألقيت قطعة نقود 20 مرة. احسب احتمال الحصول على 8 صور باستخدام:

(ا) توزيع ذي الحدين (ب) التوزيع الطبيعي كتقريب لذي الحدين .

-١٠٥- إذا كان % 70 من الطلاب الملتحقين بالكليات يحصلون على مؤهلاتهم. أوجد احتمــال أنه من بين 20 طالبا مختارين عشوائيا من الملتحقين حديثا سوف يحصل على أكثر من 10 طـــلاب منهم على المؤهل (باستخدام التوزيع الطبيعي كتقريب لذي الحدين) .

-١٠٦- بفرض أن 45% من المكالمات التي يستقبلها عامل تليفون في شركة ما مـــن مســافات بعيدة. ما هو احتمال الحصول على 11 مكالمة من مسافة بعيدة من بين 20 مكالمـــة يســتقبلهم ، استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين.

-١٠٧- تستهلك شركة ما كمية من الثلج يوميا في المتوسط 600 رطل بانحراف معيـــاري 25 رطل. فإذا كانت الكمية المستهلكة يوميا تتبع التوزيع الطبيعي. أوجد (ا) احتمال أن تســـتهلك الشركة اكثر من 700 رطل يوميا (ب) احتمال أن تستهلك من 500 إلى 800 يوميا.

-١٠٨- إذا كانت $p=0.3$ استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين لتقدير احتمـــال الحصول على 140 وحدة تالفة من بين 500 وحدة تالفة منتجة من إحدى المصانع.

-١٠٩- إذا كان معروف أن نقطة الذوبان للذهب هي $1.06c^{\circ}$ (في المتوسط) بانحراف معيـاري $0.3\ c^{\circ}$ أوجد احتمال $P(X>1.77)$.

-١١٠- إذا كان الطلب على اللحوم في مخزن لبيع اللحوم، خلال أسبوع ، تقريبا يتبع التوزيـــع الطبيعي بمتوسط 5000 رطل والحراف معياري 300 رطل أوجدي $P(X>5300)$ في اسبوع.

-١١١- إذا كان احتمال أن تنجب سيدة مولود ذكر 0.5 . فإذا كان هناك 100 سيدة في هــذه القرية حوامل وبافتراض أن X ترمز لعدد المواليد الذكور أوجد (ا) التوزيــع الاحتمــالي للمتغــير العشوائي X (ب) احتمال أن تنجب على الأقل 3 مواليد ذكور (جـ) احتمال عدم إنجاب طفــل ذكر (استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين) .

-١١٢- صممت سيارة جديدة تحت فرض أن 70% من المبيعات سوف تكون للســيدات . إذا اختيرت عينة عشوائية من 500 مشترى . مــا هــو احتمــال أنــه علــى الأقــل 270 منــهم سيدات . (استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين) .

-١١٣- إذا كان احتمال بقاء مريض لأكثر من ثمانية وأربعين ساعة فى قســم العنايــة المركــزة بمستشفى هو 0.055 ، المطلوب إيجاد احتمال أن يمكث عشرة مرضى أكثر من ثمانية وأربعين ساعة من بين خمسين مريضا الحقوا بالقسم في يوم محدد (استخدم التوزيع الطبيعي كتقريب لتوزيــع ذي الحدين) .

الفصل السابع

توزيعات المعاينة

Sampling Distributions

(٧-١) مقدمـــــة **Introduction**

يهتم فرع الإحصاء الاستدلالى بالتعميم والتنبؤ، فعلى سبيل المثال يمكـــن القـــول أن متوسط دخل الفرد في بلد ما $ 86000 في السنة وذلك بناء على عينة عشوائية اختـــيرت من هذا البلد. وبتوضيح آخر يمكن أن نتوقع بناء على أراء مجموعـــة مـــن الأشـــخاص في الشارع أن 80% من أصوات الناخبين في مدينة ما سوف تعطى لمرشح معين. كما يمكـــن التوقع أن عمر المصباح الكهربائى من إنتاج مصنع ما يتراوح بين 1150 ســـاعة و 1250 ساعة بدرجة ثقة معينة. فإننا نجد فى كل مثال من الأمثلة السابقة، تم حساب إحصـــاء مـــن عينة عشوائية تم اختيارها من المجتمع موضع الدراسة، ومن تلك الإحصاءات أمكننا الوصول إلى جمل تخص قيم المعالم والتى قد تكون صحيحة أو غير صحيحة. التعميم من الإحصـــاء إلى المعلمة يكون بثقة فقط إذا استطعنا أن نفهم العلاقة بين المجتمع وعيناته.

في الفصل الخامس عرفنا الإحصاء على أنه متغير عشوائى يعتمد قيمته فقط على العينة، وبالتالي فإن نفس الحسابات لعينات مختلفة من المجتمع تؤدى إلى قيم مختلفة للإحصاء. هـــذه الاختلافات في قيم الإحصاء تعتمد على حجم المجتمع وحجـــم العينـــات والطريقـــة الـــتى استخدمت في اختيار العينات العشوائية. إذا كان حجم المجتمع كبيراً أو لانهـــائي فـــإن التوزيع الاحتمالى للإحصاء في حالة السحب بإرجاع سوف يكون نفسه في حالة الســـحب بدون إرجاع. ومن ناحية أخري فإن السحب بإرجاع من مجتمع صغير محدود يعطى توزيعـــاً للإحصاء يختلف قليلاً عن السحب بدون إرجاع. أخيرا المعاينة مع الإرجاع مـــن مجتمـــع محدود يكافئ المعاينة من مجتمع لانهائى وذلك لعدم وجود حدود لحجم العينة المختارة مـــن المجتمع.

تعريف : التوزيـــع الاحتمـــالي لأي إحصـــاء يســـمى التوزيـــع العيـــنى **sampling distribution**

تعريف : الانحراف المعياري للتوزيع العينى لأى إحصـــاء يســـمى الخطـــأ المعيـــارى **standard error** للإحصاء .

فعلى سبيل المثال التوزيع الاحتمالى للإحصاء $\overline{X}$ يسمى التوزيع العينى للمتوسط، كمـــا أن الخطأ المعيارى للمتوسط هو الانحراف المعيارى للتوزيع العينى للإحصاء $\overline{X}$.

في هذا الفصل سوف ندرس بعض توزيعات المعاينة الأكثر اســـتخداما في الإحصـــاء. التطبيقات على تلك التوزيعات العينية تخص مشاكل الاستدلال الإحصـــائى الـــتى ســـوف نتناولها في الفصل الثامن والتاسع.

(٧-٢) توزيعات المعاينة الطبيعية Normal Sampling Distributions

إذا أخذنا عينات متكررة من الحجم n من توزيع متصل له متوسط μ وبتبـــاين σ^2. لكل عينة ثم حساب القيمة y لإحصاء ما Y، والذي نفسه متغير عشوائي متصل. بفـــرض أن Y يتبع توزيعاً طبيعياً بمتوسط μ_Y وانحراف معياري σ_Y. النظرية التالية تنص على أن :

نظرية (٧-١) إذا كان y قيمة للإحصاءY والذي يتبع توزيعاً طبيعياً بمتوسط μ_Yوانحراف معياري σ_Y، فإن :

$$z = \frac{y - \mu_Y}{\sigma_Y}$$

هي قيمة لمتغير عشوائى Z يتبع التوزيع الطبيعي القياسى حيث :

$$Z = \frac{Y - \mu_Y}{\sigma_Y}.$$

يمكن تطبيق نظرية (٧-١) للإحصاءات المحسوبة من عينات عشـــوائية اختـــيرت مـــن مجتمعات متقطعة، سواء محدودة أو غير محدودة، والتي التوزيعات العينية لإحصاءاتها تقريبـــاً تتبع توزيعاً طبيعياً. هذا ويمكننا استخدام جدول التوزيع الطبيعي في الملحق (٣) في حســـاب الاحتمالات التي يأخذها الإحصاء في فترات معينة.

(٧-٣) توزيعات المعاينة للمتوسط Sampling Distributions of the Mean

يعتبر توزيع المعاينة للمتوسط $\overline{X}$ أهم توزيع معاينة سوف نتناوله في هذا الفصـــل. إن شكل ونوع التوزيع الاحتمالى لمجتمع متوسط العينات (التوزيع العيني للمتوســـط) يعتمـــد على شكل المجتمع الأصلى الذى اختيرت منه العينات. النظرية الآتية تعطي التوزيع العيـــنى للمتوسط إذا كان المجتمع الأصلي التي اختيرت منه العينات يتبع توزيعاً طبيعيـــاً ونذكرهـــا بدون برهان.

نظرية (٧-٢) إذا أخذنا عينات متكررة من مجتمع معروف أنه يتبع توزيعـــاً طبيعيـــاً بمتوسط μ و انحراف معياري σ فإن التوزيع العينى للإحصاء $\overline{X}$ يتبع توزيعاً طبيعياً بمتوسط $\mu_{\bar{x}} = \mu$ وانحراف معياري $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ حيث $\mu_{\bar{x}}$ و $\sigma_{\bar{x}}$ يرمزان للمتوسط والانحراف المعياري على التوالي للتوزيع العينى للإحصاء $\overline{X}$.

مثال (٧-١) إذا كانت أوزان الطلاب في جامعة ما تتبع توزيعاً طبيعياً متوسطة 70 كيلـــو جراما وانحرافه المعياري 10 كيلو جراما. اختيرت عينه عشوائية مكونة من 25 طالبا فما هو

احتمال أن يكون متوسط الأوزان أقل من 75 كيلو جرام.

الحل .

$$\mu = 70 \quad , \quad \sigma = 10 \quad , \quad n = 25$$

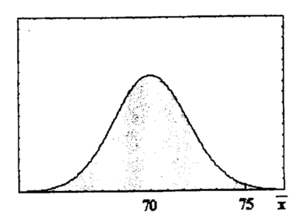
تبعا لنظرية (٧-١) فإن $\overline{X}$ يتبع توزيعاً طبيعياً متوسطة $\mu_{\overline{X}} = \mu = 70$ وانحرافه المعياري :

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2.$$

والمطلوب هو حساب الاحتمال :

$$P(\overline{X} < 75) .$$

والذي يساوى المساحة المظللة في شكل (٧-١).



شكل (٧-١)

عندما $\overline{x} = 75$ فإن قيمة z المقابلة لها هي :

$$z = \frac{\overline{x} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{75 - 70}{2} = 2.5$$

وعلى ذلك :

$$\begin{aligned} P(\overline{X} < 75) &= P(Z < 2.5) \\ &= 1\text{-}P(Z > 2.5) \\ &= 0.5 + P(0 < Z < 2.5) \\ &= 0.5 + 0.4938 = 0.9938 . \end{aligned}$$

في كثير من الحالات يكون التوزيع الاحتمالى للمجتمع الأصلي غير طبيعى ويتطلب الأمـــر معرفة التوزيع الاحتمالى للوسط الحســـابي $\overline{X}$. للتسهيل سوف نستخدم مجتمعا منفصـــلا

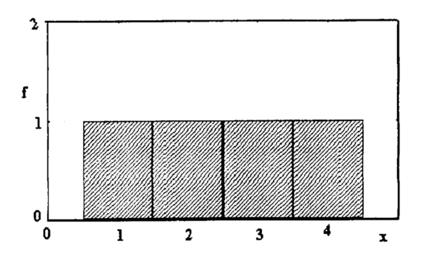
منتظما يتكون من القيم 1, 2, 3, 4 والذى متوسطة :

$$\mu = \frac{1+2+3+4}{4} = 2.5$$

وانحرافه المعياري :

$$\sigma = \sqrt{\frac{(1-2.5)^2 + (2-2.5)^2 + (3-2.5)^2 + (4-2.5)^2}{4}} = 1.118.$$

المدرج التكراري لهذا المجتمع موضح في شكل (٧-٢).



شكل (٧-٢)

بفرض أنه تم اختيار كل العينات من الحجم n = 2 من هذا المجتمع بإرجاع والذى يكافئ المعاينة من مجتمع لانهائي. يعطى جدول (٧-١) كل العينات الممكنة التى يمكن اختيارها (عدد العينات $N^n = 4^2 = 16$ عينة) من هذا المجتمع مع قيمها. لكل عينه تم حساب $\bar{x}$ والتوزيع التكراري لمجتمع متوسط العينات التى حجم كل منها n= 2 معطى في جدول (٧-٢) وتوزيعه التكراري في شكل (٧-٣).

جدول (٧-١)

رقم العينة	1	2	3	4	5	6	7	8
القيم	1,1	1,2	1,3	1,4	2,1	2,2	2,3	2,4
رقم العينة	9	10	11	12	13	14	15	16
القيم	3,1	3,2	3,3	3,4	4,1	4,2	4,3	4,4

جدول (٧-٢)

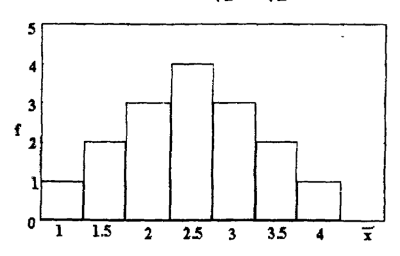
$\overline{x}$	1	1.5	2	2.5	3	3.5	4
f التكرار	1	2	3	4	3	2	1

يلاحظ أن توزيع المعاينة للإحصاء $\overline{X}$ في شكل (٧ -٣) تقريبا طبيعي. المتوسط و الانحراف المعياري للتوزيع العيني للإحصاء $\overline{X}$ تم حسابهما من جدول (٧ -٢) وهما على التوالي :

$$\mu_{\overline{X}} = \frac{\Sigma f\,\overline{x}}{\Sigma f} = \frac{40}{16} = 2.5 = \mu \ ,$$

$$\sigma_{\overline{X}} = \sqrt{\frac{\Sigma f(\overline{x}-2.5)^2}{\Sigma f}} = \sqrt{\frac{10}{16}}$$

$$= 0.791 = \frac{1.118}{\sqrt{2}} = \frac{\sigma}{\sqrt{n}}.$$



شكل (٧-٣)

دائما المتوسط للإحصاء $\overline{X}$ يساوى متوسط المجتمع الذى اختيرت منه العينات العشوائية ولا يعتمد على حجم العينة. بينما الانحراف المعيارى للإحصاء $\overline{X}$ يعتمد على حجم العينة ويساوى الانحراف المعياري للمجتمع الأصلي σ مقسوما على $\sqrt{n}$. وعلى ذلك كلما زادت حجم العينة كلما قل الخطأ المعيارى للإحصاء $\overline{X}$ واقتربت $\overline{x}$ من μ وعلى ذلك يمكننا استخدام $\overline{x}$ كتقدير للمعلمة μ.

نظرية (٧ - ٣) إذا اختيرت كل العينات الممكنة من الحجم n بإرجاع من مجتمع محدود

من الحجم N وله متوسط μ وانحراف معياري σ فإن التوزيع العينى للإحصاء $\overline{X}$ تقريبـــاً يتبع توزيعاً طبيعياً بمتوسط $\mu_{\overline{X}} = \mu$ وانحراف معياري $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ وعلى ذلك :

$$z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائى Z يتبع التوزيع الطبيعى القياسى . نظرية (٧-٣) صحيحـــة لأي مجتمع محدود عندما $n \geq 30$.

مثال (٧ - ٢) مجتمع مكون من المفردات الآتية :

2 , 2 , 2 , 4 , 5 , 7 , 7 , 7, 8

أوجد احتمال أن عينة عشوائية من الحجم $n = 35$ اختيرت من هذا المجتمع بإرجاع تعطـــى متوسط عينة أكبر من 5.

الحل . المتوسط والانحراف المعياري للمجتمع المعطى هما :

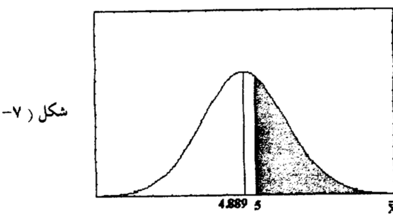
$$\mu = 4.889 \quad , \quad \sigma = 2.331 .$$

وحيث أن $n>30$ وتبعا لنظرية (٧-٣) فإن التوزيع العيني للإحصـــاء $\overline{X}$ تقريبـــاً يتبـــع التوزيـــــع الطبيعـــــــي بمتوســـــــط $\mu_{\overline{X}} = \mu = 4.889$ وانحـــــراف معياري $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2.331}{\sqrt{35}} = 0.394$. احتمال أن $\overline{X}$ أكبر من 5 يساوى المساحة المظللة في شكل (٧-٤). قيمة z المقابلة لقيمة $\overline{x} = 5$ هي :

$$z = \frac{\overline{x} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{5 - 4.889}{0.394} = 0.282.$$

وعلى ذلك :

$$P(\overline{X} > 5) = P(Z > 0.282)$$
$$= 0.5 - P(0 < Z < 0.282) .$$
$$= 0.5 - 0.1103 = 0.3897.$$



شكل (٧-٤)

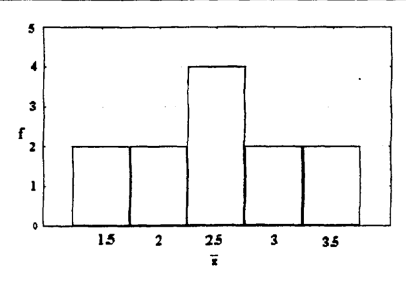
بفرض أننا سحبنا كل العينات الممكنة من الحجم $n = 2$ من مجتمعنا المنتظـــم والــذى مشاهداته 1, 2, 3, 4 ولكن بدون إرجاع. لكل عينة تم حساب متوسط العينة $\bar{x}$. يعطـــى جدول (٧-٣) كل العينات الممكنة التى يمكن اختيارها من هذا المجتمع بدون إرجاع مـــع قيم كل عينة (عدد العينات $\frac{N!}{(N-n)!} = \frac{4!}{2!} = 12$ عينة). التوزيع التكـــراري لمجتمـــع متوسط العينات من الحجم $n = 2$ معطى في جدول (٧-٤). المدرج التكراري لمجتمـــع متوسط العينات موضح في شكل (٧-٥).

جدول (٧-٣)

رقم العينة	1	2	3	4	5	6
القيم	1,2	1,3	1,4	2,1	2,3	2,4
رقم العينة	7	8	9	10	11	12
القيم	3,1	3,2	3,4	4,1	4,2	4,3

جدول (٧ - ٤)

$\bar{x}$	1.5	2.0	2.5	3.0	3.5
f	2	2	4	2	2



شكل (٧-٥)

يتضح من شكل (٧-٥) أن التوزيع العينى للإحصاء $\bar{X}$ في حالة السحب بدون إرجـــاع من مجتمع محدود بعيدا عن التوزيع الطبيعى حيث $n = 2$. من جـــدول (٧-٤) يمكـــن

حساب المتوسط والانحراف المعياري للإحصاء $\overline{X}$ على التوالي :

$$\mu_{\overline{X}} = \frac{\Sigma f \bar{x}}{\Sigma f} = \frac{30}{12} = 2.5 = \mu \quad ,$$

$$\sigma_{\overline{X}} = \sqrt{\frac{\Sigma f (\bar{x} - 2.5)^2}{\Sigma f}} = \sqrt{\frac{5}{12}} = 0.645$$

$$= \frac{1.118}{\sqrt{2}} \sqrt{\frac{4-2}{4-1}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = 0.645 .$$

عندما يكون $n \geq 30$ يمكن تطبيق النظرية التالية :

نظرية (٧-٤) إذا اختيرت كل العينات الممكنة بدون إرجاع من مجتمع محدود من الحجم N وله متوسط μ وانحراف معياري σ، فإن التوزيع العينى للإحصاء $\overline{X}$ تقريبا يتبع توزيعـــا طبيعيا بمتوسط وانحراف معياري معطى كالتالي :

$$\mu_{\overline{X}} = \mu,$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}.$$

يسمى المقدار $\sqrt{\frac{N-n}{N-1}}$ معامل التصحيح. إذا كان حجم العينة صغيرا جدا بالنسبة لحجم المجتمع فإن $(N-n) / (N-1)$ تكون قريبة من 1 ويمكن إسقاطها من المعادلة. وقـــد جـــرت العادة على إهمال هذا الحد عندما تكون $n < 0.05\, N$.

مثال (٧-٣) من مجتمع مكون من القيم

2, 2, 2, 4, 5, 6, 7, 7, 7, 8

المطلوب : (أ) حساب متوسط المجتمع الأصلى μ وانحرافه المعيارى σ.

(ب) حساب متوسط مجتمع متوسطات العينات $\mu_{\overline{X}}$ وانحرافـــه المعيـــاري $\sigma_{\overline{X}}$ عند n=2 (السحب بدون إرجاع).

الحل .

(أ) $\mu = 5, \ \sigma = 2.24 , \ N = 10, \ n = 2$

(ب) $\mu_{\overline{X}} = \mu = 5$

وحيث أن :

$$0.05\, N = (0.05)\,(10) = 0.5$$

أى أن $n = 2$ سوف تكون أكبر من 0.5 $(n > .05\, N)$ وعلى ذلك لا يمكننا إهمال معامل التصحيح في صيغة $\sigma_{\overline{X}}$. وعلى ذلك :

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{2.24}{\sqrt{2}}\sqrt{\frac{10-2}{10-1}} = 1.4933.$$

مثال (٧-٤) افترض مجتمعا ما يتكون من 1000 عنصر له متوسط حسابى $\mu=15$ وانحراف معياري $\sigma = 6$ أوجد :

(أ) المتوسط الحسابى والانحراف المعيارى لتوزيع المعاينة للإحصاء $\overline{X}$ عندما يكون $n = 36$ (السحب بدون إرجاع).

(ب) احتمال أن يقع متوسط العينة العشوائية من الحجم $n = 36$ بين 13 ,16.

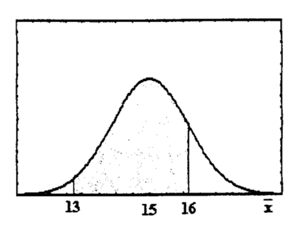
الحل .

(أ) بما أن $n = 36 < 0.05\, N = 50$ فإن التوزيع العينى للإحصاء $\overline{X}$ يكون له متوسط وانحراف معيارى كالتالي :

$$\mu_{\overline{X}} = \mu = 15,$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{36}} = 1.$$

(ب) الاحتمال المطلوب موضح بالمساحة المظللة في شكل (٧-٦) .



شكل (٧-٦)

قيمة z_1 المقابلة لقيمة $\overline{x}_1=13$ هي :

$$z_1 = \frac{\overline{x}_1 - \mu}{\sigma_{\overline{x}}} = \frac{13-15}{1} = -2,$$

قيمة z_2 المقابلة للقيمة $\overline{x}_2 = 16$ هي :

$$z_2 = \frac{\overline{x}_2 - \mu}{\sigma_{\overline{x}}} = \frac{16-15}{1} = 1.$$

وعلى ذلك :

$$P(13 < \overline{X} < 16) = P(-2 < Z < 1)$$
$$= P0 < Z < 2) + P(0 < Z < 1)$$
$$= 0.4772 + 0.3413 = 0.8185.$$

للمجتمعات الكبيرة أو اللانهائية سواء كانت متصلة أم متقطعة تنص النظرية التالية على .

نظرية (٧–٥) إذا اختيرت كل العينات الممكنة من الحجم n من مجتمع كبير أو لانهائي بمتوسط μ وتباين σ^2 فإن التوزيع العينى للإحصاء $\overline{X}$ تقريبا يتبع توزيعا طبيعيا بمتوسط :

$$\mu_{\overline{X}} = \mu$$

وانحراف معياري :

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}.$$

وعلى ذلك فإن :

$$z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائى Z يتبع التوزيع الطبيعى القياسي٠

التقريب الطبيعي في نظرية (٧–٥) سوف يكون جيدا إذا كانت $n \geq 30$ بصرف النظر عن شكل المجتمع الأصلى الذى اختيرت منه العينات. إذا كانت $n<30$ التقريب سوف يكون جيد فقط إذا كان المجتمع لا يختلف كثيرا عن التوزيع الطبيعي٠

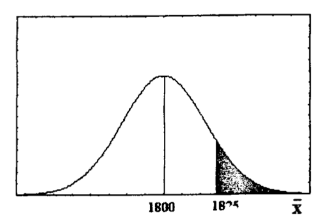
مثال (٧–٥) إذا كانت أعمار المصابيح المنتجة بواسطة أحد المصانع لها متوسط عمر $\mu = 1800$ ساعة بانحراف معيارى $\sigma = 200$ ساعة. أوجد احتمال أن عينة عشوائية من 100 مصباح سوف يكون لها متوسط عمر أكبر من 1825 ساعة.

الحل .

المجتمع كبير والعينة هنا كبيرة n=100، على ذلك التوزيع العينى للإحصاء $\overline{X}$ تقريبا يتبع توزيعا طبيعيا بمتوسط 1800 ساعة وانحراف معياري $20 = \frac{200}{\sqrt{100}} = \frac{\sigma}{\sqrt{n}}$. وعلى ذلك عندما يكون $\overline{x} = 1825$ فإن :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{1825 - 1800}{20} = 1.25.$$

الاحتمال المطلوب موضح بالمساحة المظللة في شكل (٧-٧).



شكل (٧-٧)

وعلى ذلك :

$$P(\overline{X} > 1825) = P(Z > 1.25)$$
$$= 0.5 - 0.3944 = 0.1056.$$

مثال (٧-٦) إذا كان متوسط الدب الأسترالي $\mu = 20$ بوصة بانحراف معياري $\sigma = 4$ بوصة. فإذا خطط لاختيار عينة عشوائية من الحجم $n = 64$. ما هو احتمال أن يكون متوسط العينة أكبر من **21** بوصة.

الحل . التوزيع العيني للإحصاء $\overline{X}$ تقريبا يتبع توزيعا طبيعيا بمتوسط $\mu = 20$ وانحراف معياري :

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{64}} = 0.5.$$

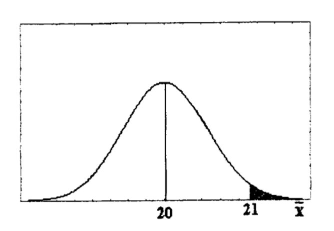
احتمال أن يكون متوسط عينه عشوائية من الحجم $n = 64$ أكبر من **21** بوصة يساوى المساحة المظللة في شكل (٧-٨).

عندما يكون $\bar{x} = 21$ فإن :

$$z = \frac{\bar{x} - \mu}{\sigma_{\overline{X}}} = \frac{21 - 20}{0.5} = 2.$$

وعلى ذلك :

$$P(\overline{X} > 21) = P(Z > 2) = 0.5 - P(0 < Z < 2)$$
$$= 0.5 - 0.4772 = 0.0228.$$



شكل (٧-٨)

مثال (٧-٧) إذا كان متوسط سمك الرخام المنتج في مصنع للعب الأطفال هو 0.85سم بانحراف معياري 0.01 المطلوب :

(أ) احتمال اختيار عينة عشوائية من 100 قطعة رخام لها سمك أكبر من 0.851.

(ب) ما هما القيمتان التى تتوقع أن يكون 95% من متوسطات العينات بينهما .

الحل .

(أ) التوزيع العينى للإحصاء $\overline{X}$ تقريبا يتبع توزيعا طبيعيا بمتوسط $\mu_{\overline{X}} = \mu = 0.85.$ وانحراف معياري :

$$\frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{100}} = 0.001.$$

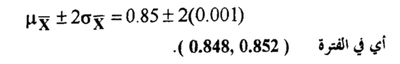
عندما يكون $\overline{x} = 0.851$ فإن :

$$z = \frac{\overline{x} - \mu}{\sigma_{\overline{X}}} = \frac{0.851 - 0.85}{0.001} = 1.0.$$

الاحتمال المطلوب يساوي المساحة المظللة في شكل (٧-٩) وعلى ذلك :

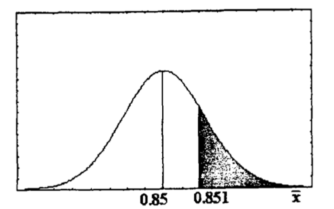
$$P(\overline{X} > 0.851) = P(Z > 1.0)$$
$$= 0.5 - 0.3413 = 0.1587.$$

(ب) باستخدام القاعدة التجريبية فإننا نتوقع أن % 95 متوسطات العينات تقع في الفترة :



$$\mu_{\overline{X}} \pm 2\sigma_{\overline{X}} = 0.85 \pm 2(0.001)$$

أي في الفترة (0.848, 0.852).



شكل (٧-٩)

(٧-٤) التوزيعات العينية للفرق بين متوسطى مجتمعين

Sampling Distributions of the Different Between Two Populations Means

بفرض أن لدينا مجتمعين الأول متوسطة μ_1 وتباينه σ_1^2 والثاني متوسطة μ_2 وتباينه σ_2^2. بفرض أن قيم المتغير $\overline{X}_1$ تمثل متوسطات لعينات عشوائية من الحجم n_1 اختيرت من المجتمع الأول، وقيم المتغير $\overline{X}_2$ تمثل متوسطات لعينات عشوائية من المجتمع الثاني ومستقلة عن المجتمع الأول. التوزيع للفروق $\bar{x}_1 - \bar{x}_2$ بين الفئتين من متوسطات العينتين المستقلتين يسمى التوزيع العينى للإحصاء $\overline{X}_1 - \overline{X}_2$. للتسهيل نفرض أن المجتمع الأول من الحجم $N_1 = 3$ ويتكون من القيم 4, 5, 6 والتى متوسطها :

$$\mu_1 = \frac{4+5+6}{3} = 5$$

وتباينها :

$$\sigma_1^2 = \frac{(4-5)^2 + (5-5)^2 + (6-5)^2}{3} = \frac{2}{3}.$$

المجتمع الثاني يتكون من القيمتين 1, 4 ولهما المتوسط :

$$\mu_2 = \frac{1+4}{2} = 2.5$$

والتباين :

$$\sigma_2^2 = \frac{(1-2.5)^2 + (4-2.5)^2}{2} = \frac{9}{4}.$$

من المجتمع الأول تم اختيار كل العينات الممكنة من الحجم $n_1 = 2$ مع الإرجاع وحساب المتوسط $\bar{x}$ لكل عينة. بنفس الشكل للمجتمع الثاني تم اختيار كل العينات الممكنة من الحجم $n_2 = 3$ وحساب $\bar{x}_2$ لكل عينة. الفئتان من كل العينات ومتوسطاتها معطاة في جدول (٧-٥).

جدول (٧-٥)

المجتمع الأول	رقم العينة	1	2	3	4	5	6	7	8	9
	القيم	4,4	4,5	4,6	5,4	5,5	5,6	6,4	6,5	6,6
	$\bar{x}_1$	4.0	4.5	5.0	4.5	5.0	5.5	5.0	5.5	6.0
المجتمع الثاني	رقم العينة	1	2	3	4	5	6	7	8	
	القيم	1,1,1	1,1,4	1,4,1	4,1,1	4,4,1	1,4,4	4,1,4	4,4,4	
	$\bar{x}_2$	1	2	2	2	3	3	3	4	

الفروق الممكنة والتي عددها 72 من $\bar{x}_1 - \bar{x}_2$ معطاة في جدول (٧-٦).

جدول (٧-٦)

$\bar{x}_2$	$\bar{x}_1$								
	4.0	4.5	5.0	4.5	5.0	5.5	5.0	5.5	6.0
1	3.0	3.5	4.0	3.5	4.0	4.5	4.0	4.5	5.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
4	0.0	0.5	1.0	0.5	1.0	1.5	1.0	1.5	2.0

التوزيع التكرارى للإحصاء $\bar{X}_1 - \bar{X}_2$ في جدول (٧-٧) ومدرجة التكرارى في شكل (٧-١٠). من الواضح أن المتغير العشوائى $\bar{X}_1 - \bar{X}_2$ تقريبا يتبع توزيعا طبيعيا وهذا

التقريب يتحسن عندما يزيد n_1, n_2. بتطبيق نظرية (٤-٦) ثم نظريـــة (٧-٣) فـــإن متوسط الفروق لمجتمع العينات المستقلة هو :

$$\mu_{\overline{X}_1-\overline{X}_2} = E(\overline{X}_1 - \overline{X}_2) = E(\overline{X}_1) - E(\overline{X}_2) = \mu_1 - \mu_2.$$

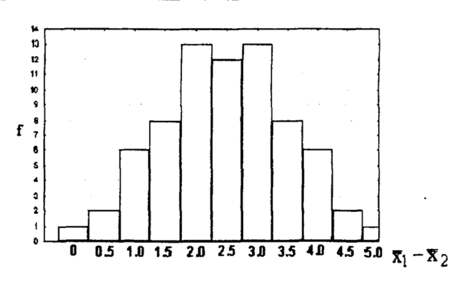
هذه النتيجة يمكن التحقق منها من البيانات في جدول (٧-٧) حيث :

$$\mu_{\overline{X}_1-\overline{X}_2} = \frac{\Sigma f(\bar{x}_1 - \bar{x}_2)}{\Sigma\ f} = \frac{180}{72} = 2.5 = 5 - 2.5$$

$$= \mu_1 - \mu_2.$$

جدول (٧-٧)

$\bar{x}_1 - \bar{x}_2$	0.0	.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
f	1	2	6	8	13	12	13	8	6	2	1



شكل (٧-١٠)

أيضا بتطبيق نظرية (٤-٨) ثم نظرية (٧-٣) فإن التباين لفروق المتوسطات المستقلة هو :

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \sigma^2_{\bar{X}_1} + \sigma^2_{\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$= \left[\left(\frac{2}{3}\right)/2\right] + \left[\left(\frac{9}{4}\right)/3\right] = 1.08333.$$

هذه النتيجة يمكن التحقق منها بسهولة وذلك بحساب التباين (1.08333) من البيانات في جدول (٧-٧). النتائج التي تم الحصول عليها للتوزيع العيني للإحصاء $\overline{X}_1 - \overline{X}_2$ عند المعاينة بإرجاع من مجتمع محدود تكون صحيحة للمجتمعات اللانهائية سواء المتقطعة أو المتصلة وأيضا للمجتمعات المحدودة عند المعاينة بدون إرجاع بشرط أن أحجام المجتمعات N_2 , N_1 تكون كبيرة نسبيا عن أحجام العينات n_2 ، n_1 على التوالي. أما إذا كان حجم المجتمع صغيرا والسحب بدون إرجاع فلابد من حساب $\sigma^2_{\overline{X}_1}, \sigma^2_{\overline{X}_2}$ من صيغة $\sigma_{\overline{X}}$ في نظرية (٧-٤).

تقتصر الدراسة من الآن وفي الفصول التالية على التوزيع العيني للفروق بين المتوسطات المستقلة فقط إذا كان حجم المجتمع الذى تختار منه العينات كبيراً أو لانهائي.

نظرية (٧-٦) إذا اختيرت عينات مستقلة من الحجم N_2 ، N_1 من مجتمعين كبيرين (أو لانهائيين) ، متقطعة أو متصلة، بمتوسطي μ_1, μ_2 وتبايني σ_1^2, σ_2^2 على التوالي، فإن التوزيع العينى لفروق المتوسطات، $\overline{X}_1 - \overline{X}_2$ ، تقريباً يتبع توزيعاً طبيعياً بمتوسط وانحراف معياري معطى كالتالي :

$$\mu_{\overline{X}_1 - \overline{X}_2} = \mu_1 - \mu_2,$$

$$\sigma_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

وعلى ذلك :

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

هي قيمة لمتغير عشوائى Z يتبع التوزيع الطبيعى القياسى.

إذا كان كل من n_2 ، n_1 أكبر من أو يساوى 30 ، فإن التقريب الطبيعى لتوزيع $\overline{X}_1 - \overline{X}_2$ يكون جيداً جداً.

ملحوظة : إذا كانت العينتين تم اختيارهما من مجتمعين طبيعيين فإن :

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

هي قيمة لمتغير عشوائي Z يتبع التوزيع الطبيعي القياسي بصرف النظر عن حجم كلا من n_2 ، n_1 .

مثال (٧-٨) ينتج المصنع A بطاريات سيارة لها متوسط عمر 3.5 سنة بانحراف معياري 0.45 سنة. نفس البطاريات تنتج من المصنع B بمتوسط عمر 3.3 سنة وانحراف معياري 0.3 سنة. ما هو احتمال أن عينة عشوائية من 30 بطارية من المصنع A يكون لها متوسط عمر على الأقل يزيد 0.4 سنة عن متوسط عمر 36 بطارية من المصنع B ؟

الحل . سوف يكون لدينا البيانات التالية :

المجتمع الثاني	المجتمع الأول
$\mu_2 = 3.3$	$\mu_1 = 3.5$
$\sigma_2 = 0.3$	$\sigma_1 = 0.45$
$n_2 = 36$	$n_1 = 30$

العينات هنا كبيرة بدرجة كافية بحيث أن كل من $\overline{X}_2$ و $\overline{X}_1$ تقريبا يتبع توزيعا طبيعيا وعلى ذلك فإن $\overline{X}_1 - \overline{X}_2$ تقريبا تتبع توزيعا طبيعيا بمتوسط وانحراف معياري على التوالي :

$$\mu_{\overline{X}_1-\overline{X}_2} = \mu_1 - \mu_2 = 3.5 - 3.3 = 0.2,$$

$$\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

$$= \sqrt{\frac{0.45^2}{30} + \frac{0.3^2}{36}} = 0.096177,$$

الاحتمال المطلوب موضح بالمساحة المظللة في شكل (٧-١١).

عندما : $\bar{x}_1 - \bar{x}_2 = 0.4$

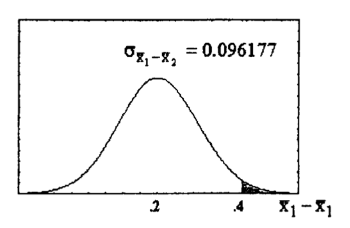
فإن :

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\overline{X}_1-\overline{X}_2}} = \frac{0.4 - 0.2}{0.096177} = 2.08 ,$$

وعلى ذلك :

$$P(\overline{X}_1 - \overline{X}_2 > 0.4) = P(Z > 2.08)$$
$$= 0.5 - P(0 < Z < 2.08)$$

$$= 0.5 - 0.4812 = 0.0188.$$



شكل (٧-١١)

(٧-٥) التوزيعات العينية للنسب

Sampling Distributions of Proportions

بفرض أن لدينا مجتمعا ما وأن بعض مفردات هذا المجتمع تتوفر فيها صفة معينة وأن نسبة هذه المفردات هي p. فعلى سبيل المثال p قد تكون نسبة الأفراد المصابين بتسوس الأسنان في مدينة ما أو نسبة الطلبة المدخنين في كلية ما أو نسبة الوحدات المعيبة في مصنع ما. .. الخ . إذا أخذنا عينة عشوائية من الحجم n من هذا المجتمع ووجدنا من بينها x مفرده تتوفر فيها الصفة، وتم حساب $\hat{p} = \frac{x}{n}$ والتي تمثل نسبة المفردات في العينة والتي تتوفر فيها الصفة المعنية. إذا أخذنا عينات متكررة من الحجم n من هذا المجتمع فإن $\hat{p}$ تتغير من عينة إلى أخرى وتمثل قيمة للإحصاء $\hat{P}$. الآن سوف نتعرف على كيفية اشتقاق التوزيع العيني للنسبة وخصائصه سواء في حالة السحب بإرجاع أو بدون إرجاع وذلك من الأمثلة التالية:

مثال (٧-٩) مجتمع يتكون من القيم 1, 2, 3, 4 فإذا تم سحب كل العينات الممكنة من الحجم n = 2 من هذا المجتمع (بإرجاع). المطلوب إيجاد التوزيع العينى للإحصاء $\hat{P}$ والذي يمثل نسبة ظهور الرقم 4 في العينة. وإثبات أن المتوسط والتباين للتوزيع العينى للإحصاء $\hat{P}$ هما على التوالي :

$$\mu_{\hat{P}} = p,$$

$$\sigma^2_{\hat{P}} = \frac{pq}{n}.$$

الحل . الجدول (٧-٨) يحتوى على كل العينات الممكنة ونسبة ظهور الرقم 4 فيها.

جدول (٧-٨)

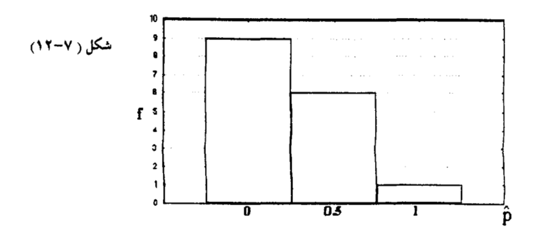
نسبة ظهور 4	عدد مرات ظهور 4	القيم	رقم العينة	نسبة ظهور 4	عدد مرات ظهور 4	القيم	رقم العينة
0	0	3,1	9	0	0	1,1	1
0	0	3,2	10	0	0	1,2	2
0	0	3,3	11	0	0	1,3	3
.5	1	3,4	12	0.5	1	1,4	4
.5	1	4,1	13	0	0	2,1	5
.5	1	4,2	14	0	0	2,2	6
.5	1	4,3	15	0	0	2,3	7
1	2	4,4	16	0.5	1	2,4	8

التوزيع التكراري لنسبة ظهور الرقم 4 للعينات من الحجم $n = 2$ التي تم اختيارها من المجتمع الذى حجمه $N = 4$ (بإرجاع) معطاة في جدول (٧-٩) ومدرجها التكراري في شكل (٧-١٢).

جدول (٧ -٩)

$\hat{p}$	0	0.5	1
f	9	6	1

التوزيع التكراري في شكل (٧ -١٢) ملتو ناحية اليمين وذلك لأن $p < 0.5$. إذا كانت $p > 0.5$ فإن التوزيع سوف يكون ملتويا ناحية اليسار .



شكل (٧-١٢)

المتوسط والتباين للتوزيع العيني للإحصاء $\hat{P}$ تم حسابهما من جدول (٧-٩) وهما :

$$\mu_{\hat{p}} = \frac{\Sigma f\,\hat{p}}{\Sigma f} = \frac{4}{16} = 0.25 = p,$$

$$\sigma^2_{\hat{p}} = \frac{\Sigma\, f(\hat{p}-0.25)^2}{\Sigma f} = 0.094$$

$$= \frac{(0.25)(0.75)}{2} = \frac{\mathbf{pq}}{\mathbf{n}}.$$

مثال (٧-١٠) أوجد التوزيع العينى للإحصاء $\hat{P}$ للبيانات في مثال (٧-٩) إذا كـــان السحب بدون إرجاع وأثبت أن متوسط وتباين التوزيع العيني للإحصاء $\hat{P}$ هما على التوالي:

$$\mu_{\hat{p}} = \mathbf{p},$$

$$\sigma^2_{\hat{p}} = \frac{\mathbf{pq}}{\mathbf{n}}\left(\frac{\mathbf{N-n}}{\mathbf{N-1}}\right).$$

الجدول (٧-١٠) يحتوى على كل العينات ونسبة ظهور الرقم 4 فيها .

جدول (٧-١٠)

رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4	رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4
1	1,2	0	0	7	3,1	0	0
2	1,3	0	0	8	3,2	0	0
3	1,4	1	.5	9	3,4	1	.5
4	2,1	0	0	10	4,1	1	.5
5	2,3	0	0	11	4,2	1	.5
6	2,4	1	.5	12	4,3	1	.5

التوزيع التكراري لنسبة ظهور الرقم 4 للعينات من الحجم $n = 2$ والتى تم اختيارها مـــن المجتمع الذى حجمه $N = 4$ (بدون إرجاع) معطاة في جـــدول (٧-١١) ومدرجـــها التكراري في شكل (٧-١٣).

جدول (٧-١١)

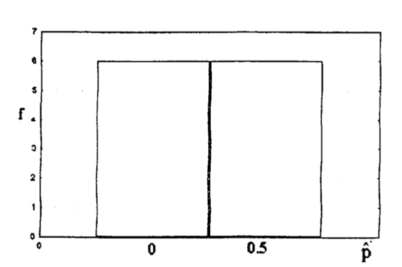
$\hat{p}$	0	0.5
f	6	6

المتوسط والتباين للتوزيع العيني للإحصاء $\hat{P}$ تم حسابهما من جدول (٧-١١) وهما :

$$\mu_{\hat{p}} = \frac{\Sigma f\,\hat{p}}{\Sigma f} = \frac{3}{12} = .25 = p,$$

$$\sigma^2_{\hat{P}} = \frac{\Sigma f(\hat{p} - 0.25)^2}{\Sigma f} = 0.0625\,.$$

$$= \frac{(0.25)(0.75)}{2}\left(\frac{4-2}{4-1}\right) = \frac{\mathbf{pq}}{\mathbf{n}}\left(\frac{\mathbf{N-n}}{\mathbf{N-1}}\right)$$



شكل (٧ -١٣)

يتضح من المثال السابق أن معامل التصحيح $\left(\frac{N-n}{N-1}\right)$ في صيغة $\sigma^2_{\hat{p}}$ يستخدم إذا كان المجتمع محدود والسحب بدون إرجاع. إذا كان حجم العينة أصغر من **0.05N** يمكن اعتبار $\left(\frac{N-n}{N-1}\right) = 1$.

مثال (٧-١١) إذا كانت نسبة المدخنين في مجتمع من الحجم **N= 350** هو **0.4**.

سحبت عينة عشوائية من الحجم $n = 60$ من هذا المجتمع (بدون إرجاع) أوجد المتوسط والتباين للتوزيع العيني للإحصاء $\hat{P}$.

الحل .

$$\mu_{\hat{P}} = .4,$$

التباين يحسب من الصيغة :

$$\sigma^2_{\hat{P}} = \frac{pq}{n}\left(\frac{N-n}{N-1}\right).$$

وذلك لأن $n=60 > 0.05N=17.1$ ، وعلى ذلك :

$$\sigma^2_{\hat{P}} = \frac{pq}{n}\left(\frac{N-n}{N-1}\right) = \frac{(0.4)(0.6)}{60}\left(\frac{350-60}{350-1}\right) = 0.0033.$$

للمجتمعات الكبيرة أو اللانهائية فإن التوزيع العيني للإحصاء $\hat{P}$ تحدده النظرية التالية :

نظرية (٧-٧) إذا كانت p هى نسبة صفة معينة في مجتمع ما واختيرت من هذا المجتمع عينات كبيرة، حجم كل منها n وكان الإحصاء $\hat{P}$ يمثل نسبة وجود هذه الصفة في العينات فإن $\hat{P}$ تقريباً تتبع توزيعاً طبيعياً متوسطة وتباينه على التوالي :

$$\mu_{\hat{P}} = p,$$

$$\sigma^2_{\hat{P}} = \frac{pq}{n}.$$

وعلى ذلك :

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

هي قيمة لمتغير عشوائى Z تقريباً يتبع التوزيع الطبيعي القياسى .

وضع العالم Cochran (1963) قواعد لحجم العينة اللازم لتطبيق نظرية (٧-٧) معطاة في جدول (٧-١٢).

جدول (٧-١٢)

يستخدم التقريب الطبيعي إذا كانn عل الأقل يساوى	إذا كانت p تساوى
30	.5
50	.4 - .6
80	.3 - .7
200	.2 - .8
600	.1 - .9
1400	.05 - .95

مثال (٧-١٢) بفرض أن مجتمعا ما أفراده عدة آلاف يمثل مصنعا لإنتاج كروت المعايدة. فإذا كان 0.2 من الكروت المنتجة تالفة (ا) أوجد المتوسط والانحراف للتوزيع العيني للإحصاء $\hat{P}$ وذلك عندما يكون $n = 300$ (ب) ما هو الاحتمال أن عينه عشوائية من الحجم $n = 300$ تعطى نسبة صفة أكبر من 0.19 (ج) أوجد احتمال أن يكون نسبة الكروت المنتجة التالفة تزيد عن 0.052 .

الحل .

(ا) بما أن المجتمع كبيراً و $n = 300,\ p = 0.2$ تحققان القواعد في جدول (٧-١٢)

وعلى ذلك فإن التوزيع العينى للإحصاء $\hat{P}$ تقريباً يتبع توزيعاً طبيعياً بمتوسط:

$$\mu_{\hat{p}} = p = 0.2$$

وانحراف معياري :

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.2)(0.8)}{300}} = 0.023.$$

(ب) عندما يكون $\hat{p} = 0.19$ فإن

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0.19 - 0.2}{0.023} = -0.43.$$

وعلى ذلك فإن :

$$\begin{aligned} P(\hat{P} > 0.19) &= P(Z > -0.43) \\ &= 0.5 + p\,(0 < Z < 0.43) \\ &= 0.5 + 0.1664 = 0.6664. \end{aligned}$$

(ج) عندما يكون $\hat{p} = 0.052$ فإن :

$$z = \frac{0.052 - 0.2}{0.023} = -6.43\,,$$

$$P(\hat{P} > 0.052) = P(Z > -6.43) \approx 1$$

مثال (٧-١٣) للمثال (٧-١١) أوجد احتمال أن تكون نسبة المدخنين في العينة أكبر من 0.35 .

الحل . عندما يكون $\hat{p} = 0.35$ فإن :

$$z = \frac{0.35 - 0.4}{0.058} = -0.86$$

وعلى ذلك فإن :

$$\begin{aligned} P(\hat{P} > 0.35) &= P(Z > -0.86) \\ &= 0.5 + P(0 < Z < 0.86) \\ &= 0.5 + 0.3051 \\ &= 0.8051. \end{aligned}$$

يمكن تطبيق نظرية (٧-٧) في حالة السحب بدون إرجاع من مجتمع محدود إذا كان حجم العينة أكبر من 0.05 N حيث متوسط التوزيع العينى للإحصاء $\hat{P}$ في هذه الحالة مازال $\mu_{\hat{P}} = p$ ولكن التباين يساوى :

$$\sigma^2_{\hat{P}} = \frac{pq}{n}\left(\frac{N-n}{N-1}\right).$$

وعلى ذلك فإن :

$$z = \frac{\hat{p} - p}{\sigma_{\hat{P}}}$$

هي قيمة لمتغير عشوائي Z يتبع التوزيع الطبيعي القياسي.

مثال (٧-١٤) إذا كانت نسبة المصابين بتسوس الأسنان في مجتمع من الحجم $N = 200$ هو 0.3. سحبت عينة عشوائية من الحجم $n = 80$ بدون إرجاع أوجد احتمال أن تكون نسبة المصابين بتسوس الأسنان في العينة أكبر من 0.35.

الحل . حجم العينة $(n=80)$ أكبر من $0.05N=10$ ، التوزيع العيني للإحصاء $\hat{P}$ سوف يكون تقريباً يتبع التوزيع الطبيعى لأن $n = 80$ تحقق القواعد في جدول (٧-١٢).
المتوسط و الانحراف المعياري للإحصاء $\hat{P}$ سوف يكون :

$$\mu_{\hat{P}} = p = 0.3\ ,$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}$$

$$= \sqrt{\frac{(.3)(.7)}{80}}\sqrt{\frac{200-80}{199}} = 0.03979.$$

عندما يكون $\hat{p} = 0.35$ فإن :

$$z = \frac{\hat{p}-p}{\sigma_{\hat{p}}} = \frac{0.35-0.3}{0.03979} \approx 1.26.$$

$$P(\hat{P} > .35) = P(Z > 1.26) = 0.5 - P(0 < Z\ 1.26)$$
$$= 0.5 - 0.3962 = 0.1038.$$

إذا كان لدينا مجتمعان مستقلان (مجتمعات كبيرة أو لانهائية) وإذا كانت p_1 هي نسبة توفر صفة ما في المجتمع الأول وكانت p_2 هى نسبة توفر الصفة نفسها في المجتمع الثانى . إذا اخترنا عينة عشوائية كبيرة حجمها n_1 من المجتمع الأول وحسبنا منها نسبة توفر الصفة محل الدراسة ولتكن $\hat{p}_1$. وإذا اخترنا عينة عشوائية كبيرة حجمها n_2 من المجتمع الثانى وحسبنا منها نسبة توفر الصفة المطلوبة ولتكن $\hat{p}_2$. بتكرار المعاينة من الحجم n_1 و n_2 فإن التوزيع العينى للإحصاء $\hat{P}_1 - \hat{P}_2$ تحدده النظرية التالية :

نظرية (٧-٨) التوزيع العينى للإحصاء $\hat{P}_1 - \hat{P}_2$ تقريباً يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{P}_1-\hat{P}_2} = p_1 - p_2,$$

وتباين :

$$\sigma^2_{\hat{P}} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}.$$

وعلى ذلك تكون :

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

هي قيمة لمتغير عشوائى Z الذى يتبع التوزيع الطبيعى القياسى .

تعطى النظرية (٧-٨) نتائجاً جيدة ، إذا كانت n_1 , n_2 محددتان طبقا لقواعد Cochran المعطاة في جدول (٧-١٢) .

(٧-٦) توزيع t   t Distribution

في معظم الأبحاث وغالبا يكون تباين المجتمع الذى تختار منه العينات مجهولا. للعينات العشوائية من الحجم $n \geq 30$ فإن التقدير الجيد للمعلمة σ^2 هو s^2. إذا كانت $n \geq 30$

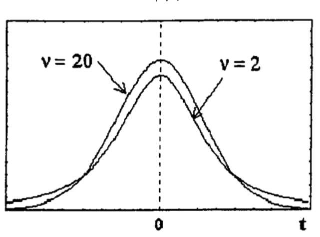
واستبدلنا σ بالقيمة S في صيغة Z لنظرية (٧-٥) فإن :

$$z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائى Z تقريبا يتبع التوزيع الطبيعى القياسى . أما إذا كان حجم العينة صغير ($n < 30$) فإن قيم $(\bar{x} - \mu)/(s/\sqrt{n})$ لا تتبع التوزيع الطبيعى القياسى . في هذه الحالة يكون اهتمامنا بتوزيع لإحصاء ما سوف نرمز له بالرمز T ، والذى قيمه تعطى من الصيغة التالية :

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

لقد تمكن ستيودنت "Student" وهو لقب لعالم إحصائي، كان ينشر أبحاثه بتوقيع ستيودنت، أن يشتق العبارة المضبوطة لتوزيع t ويسمى هذا التوزيع في كتب الإحصاء المختلفة " توزيع t" أو " توزيع ت " ، يشبه توزيع t التوزيع الطبيعى القياسى فكلاهما متماثل حول الصفر كما أن كلا التوزيعيين لهما شكل الناقوس ولكن توزيع t أكثر تشتتا وذلك راجع إلى الحقيقة أن قيم t تعتمد على الاختلاف في قيمتي $\bar{x}$ و s^2 بينما قيم z تعتمد فقط على التغير في قيمة $\bar{x}$ من عينة إلى أخرى. يختلف توزيع المتغيرT عن المتغير Z في إن التباين يعتمد على حجم العينة n ودائما أكبر من الواحد الصحيح ، فقط عندما $n \to \infty$ فإن التوزيعين يتساويان. المقام (n-1) والذي يظهر في صيغه s^2 يسمى درجات الحرية degree of freedom المرتبط بتباين العينة s^2. بتكرار المعاينة من الحجم n وحساب $\bar{x}$ و s^2 لكل عينة، فإن قيم t المقابلة يقال أنها تتبع توزيع t بدرجات حرية ν ، حيث $\nu = n-1$. وعلى ذلك سوف يكون لدينا منحنيات t مختلفة أو توزيع t لكل حجم عينه. من خصائص توزيع t أنه كلما كبرت درجات الحرية ν زاد ارتفاع منحنى t وأصبح أكثر تدببا أي اقل تشتتا وفي النهاية ينطبق على منحنى التوزيع الطبيعي القياسي.
المنحنى في شكل (٧ - ١٤) بدرجات حرية $\nu = 2$ يمثل توزيع كل قيم t المحسوبة من عينات عشوائية من الحجم $n = 3$ تكرر اختيارها من مجتمع طبيعي. بنفس الشكل، المنحنى بدرجات حرية $\nu = 20$ يمثل توزيع كل قيم t المحسوبة من عينات من الحجم $n = 21$.



شكل (٧ - ١٤)

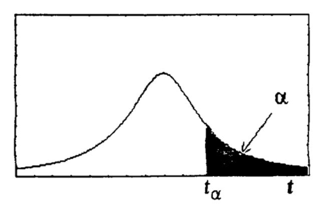
نظرية (٧-٩) إذا كان $\overline{x}$ و s^2 هما المتوسط الحسابي والتباين على التوالي لعينة عشوائية من الحجم n مأخوذة من مجتمع طبيعي له متوسط μ وتباين σ^2 غير معروف فإن :

$$t = \frac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}}.$$

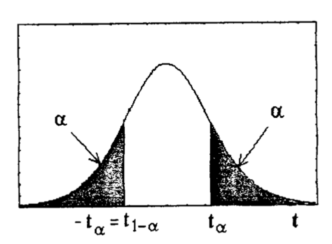
هي قيمة لمتغير عشوائي T له توزيع t بدرجات حرية $\nu = n-1$.

بفرض أن t_α ترمز لقيمة t التي توجد على المحور الأفقي تحت منحنى توزيع t بدرجات حرية ν والتي المساحة على يمينها قدرها α كما هو موضح في شكل (٧-١٥).

شكل (٧-١٥)



الجدول في ملحق (٤) يعطى قيم t_α التي تناظر الاحتمال α لدرجات حرية v حيث α تأخذ القيم التالية : 0005. ,001. ,005. ,01. ,025. ,05. ,10. ودرجات الحرية تأخذ القيم من $v = 1$ إلى $v = \infty$. يوضح الصف الثاني من الجدول قيم α والعمود الأول من الشمال قيم درجات الحرية v. أما محتويات الجدول فهى القيم t_α . ولأن المنحنى متماثل فإن $t_{1-\alpha} = -t_\alpha$ كما هو موضح في شكل (٧-١٦).



شكل (٧-١٦)

مثال (٧-١٥) أوجد (أ) قيمة $t_{.005}$, $v = 15$.

(ب) قيمة $t_{.995}$, $v = 15$.

الحل .

(أ) بالبحث في جدول توزيع t في ملحق (٤) عند تقاطع الصف $v = 15$ والعمود $\alpha = .005$ نجد أن $t_{.005} = 2.947$.

(ب) باستخدام خاصية التماثل لمنحنى توزيع t فإن $t_{.995} = -t_{.005}$ أى أن $t_{.995} = -2.947$.

مثال (٧-١٦) أوجد قيمة α حيث

$$v = 16 \quad , t_\alpha = -1.746$$

الحل .

حيث أن قيمة t سالبة فإنها تقع في الذيل الأيسر من توزيع t وباستخدام خاصية التماثل لمنحنى توزيع t فإن :

$$t_{1-\alpha} = -t_{\alpha} = 1.746$$

ومن جدول توزيع t في ملحق (٤) فإن $1-\alpha = .05$ ومنها $\alpha = .95$.

مثال (٧ -١٧) إذا كانت أعمار المصابيح الكهربائية المنتجة بواسطة أحد المصانع تتبع توزيعاً طبيعياً ، ويدعي صاحب المصنع أن متوسط أعمار هذه المصابيح هو $\mu = 500$ ساعة. لمتابعة جودة الإنتاج يأخذ 20 مصباحا كل شهر ، ويحقق الإنتاج للمواصفات القياسية إذا كانت قيمة t المحسوبة من عينة عشوائية من الحجم $n = 20$ تقع في الفترة($-t_{.05}$, $t_{.05}$) ما الاستنتاج الذي يمكن وضعه عند اختيار عينة عشوائية من الحجم $n = 20$ لها متوسط حسابي $\bar{x} = 530$ وانحراف معياري $s = 20$ ؟

الحل .

من جدول توزيع t في ملحق (٤) نجد أن $t_{.05} = 1.729$ عند درجات حرية $v = 19$.

وعلى ذلك فإن المنتج يحقق المواصفات القياسية إذا كانت قيمة t المحسوبة من عينة عشوائية من الحجم $n = 20$ مصباحا تقع في الفترة (-1.729 , 1.729) . ونجد أن :

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{530 - 500}{\frac{20}{\sqrt{20}}} = 6.708.$$

لا تقع في الفترة (-1.729 , 1.729) إذا كانت $\mu > 500$ فإن قيمة t المحسوبة من العينة تكون جيدة وتعني أن الإنتاج أفضل من المطلوب.

إذا كانت لدينا عينتان عشوائيتان مأخوذتان من مجتمعين طبيعيين بمتوسطي μ_1, μ_2 وسوف تكون نظرية (٧-٦) مفيدة فقط إذا كانت العينتان مستقلتان وتبايني المجتمعين σ_1^2, σ_2^2 معلومتان أو يمكن تقديرهما من عينات عشوائية من الحجم n_1 , n_2 حيث $n_1 \geq 30$, $n_2 \geq 30$.

إذا كانت σ_1^2, σ_2^2 مجهولتان كما يحدث في معظم الحالات ، فإن التوزيع المضبوط للمتغير Z في نظرية (٧-٦) لا يكون معروفا فيما عدا لو فرضنا أن $\sigma_1^2 = \sigma_2^2 = \sigma^2$ والتباين يمكن تقديره من الصيغة التالية :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

وعلى ذلك يكون :

$$t=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

هي قيمة للإحصاء T الذى يخضع لتوزيع t بدرجات حرية $\nu = n_1 + n_2 - 2$.

نظرية (٧-١٠) إذا كان $s_1^2, \bar{x}_1$ يمثلان المتوسط والتباين على التوالي لعينة عشوائية من الحجم n_1 مأخوذة من مجتمع طبيعي بمتوسط μ_1 وتباين مجهول σ_1^2 وإذا كانت $s_2^2, \bar{x}_2$ يمثلان المتوسط والتباين على التوالي لعينة عشوائية من الحجم n_2 مأخوذة من مجتمع طبيعى بمتوسط μ_2 وتباين σ_2^2 مجهول وإذا كانت n_1 مستقلة عن n_2 و $\sigma_1^2 = \sigma_2^2$ فإن :

$$t=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

هي قيمة لمتغير عشوائى T له توزيع t بدرجات حرية $\nu = n_1 + n_2 - 2$.

مثال (٧-١٨) يقوم مصنع بإنتاج نوعين من المصابيح الكهربائية A,B. المصابيح من النوع A لها متوسط عمر أطول 100 ساعة عن متوسط عمر المصابيح من النوع B. التباين لكلا النوعين واحد. يختار شهريا 15 مصباحا من النوع الأول، 10 مصابيح من النوع الثاني للاختبار وتحسب قيمة t. تحقق القيمة المواصفات القياسية إذا وقعت في الفترة $(-t_{0.01}\ ,\ t_{0.01})$ فإذا تم سحب في شهر ما عينة عشوائية من 15 مصباحا من المصنع A ووجد أن $s_1 = 50, \bar{x}_1 = 520$. أيضا تم سحب عينة عشوائية من المصنع B من الحجم $n_2=10$ ووجد أن $\bar{x}_2 = 500\ ,\ s_2 = 40$ هل الإنتاج يحقق المواصفات القياسية ؟

الحل . باستخدام جدول توزيع t في ملحق (٤) فإن $t_{0.01} = 2.5$ بدرجات حرية $\nu = 15 + 10 - 2 = 23$. وعلى ذلك الإنتاج يحقق المواصفات القياسية إذا كانت قيمة t المحسوبة تقع في الفترة $(-2.5\ ,\ 2.5)$. التباين المتجمع s_p^2 هو :

$$s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

$$=\frac{(14)(50)^2+(9)(40)^2}{15+10-2}=2147.826.$$

وبأخذ الجذر التربيعي للتباين المتجمع s_p^2 فإن $s_p = 46.3446$.وعلى ذلك :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{(520 - 500) - (100)}{46.3446\sqrt{\frac{1}{15} + \frac{1}{10}}} = -4.228.$$

وبما أن t لا تقع في الفترة (2.5 , -2.5) فإن الإنتاج لا يحقق المواصفات القياسية .

في بعض الأحيان بدلا من استخدام طريقة العينات المستقلة فإنه غالبا ما تستخدم طريقة العينات المتزاوجة paired samples. ففي تجارب تغذية الحيوان عند مقارنة عليقتين، حيث توضع الحيوانات المتجانسة في أزواج ويشترط أن تكون هذه الأزواج على درجة عالية من التماثل وقد تختلف الأزواج فيما بينها إلا أن أفراد كل زوج تكون متماثلة ويعطى أحد أفراد كل زوج عليقة بينما يعطى الأخر العليقة الأخرى وبالتالى فإن المقارنة بين العليقتين تتم داخل مجموعات متجانسة. في بعض الأحيان يتم ازدواج المشاهدات لمفردات العينة نفسها. فمثلا لمعرفة تأثير دواء على ارتفاع ضغط الدم نختار عينة عشوائية من الحجم n من الأشخاص ويتم قياس ضغط الدم الخاص بهم في أول فترة زمنية ثم يعالجون بهذا الدواء وبعد فترة زمنية معينة يتم قياس ضغط الدم لهم مرة أخرى . أزواج المشاهدات سوف تكون $(x_1 , y_1) , (x_2 , y_2),\ldots,(x_n, y_n)$. الفروق لأزواج المشاهدات سوف تكون $d_i = x_i - y_i , i= 1, 2, \ldots, n$. هذه الفروق تمثل قيم لمتغير عشوائى D. متوسط مجتمع الفروق μ_D سوف يساوى الصفر إذا كان الدواء ليس له تأثير. وسوف نرمز لمتوسط الفروق في عينتنا بالرمز $\bar{d}$، وسوف تختلف $\bar{d}$ من عينة إلى أخرى ولذلك يعتبر $\bar{d}$ قيمة للإحصاء $\bar{D}$. كما يحسب تباين الفروق s_d^2 والذي يعتبر قيمة للإحصاء S_d^2 لأنه يتغير من عينة إلى أخرى.

نظرية (٧ – ١١) إذا كان $d_1, d_2, \ldots, d_n$ تمثل الفروق لعدد n من أزواج المشاهدات وإذا كانت الفروق التى عددها n تمثل عينة عشوائية لها متوسط $\bar{d}$ وتباين s_d^2 مأخوذة من مجتمع الفروق الطبيعي والذى له متوسط μ_D وتباين σ_D^2 فإن :

$$t = \frac{\bar{d} - \mu_D}{\frac{s_d}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائى T له توزيع t بدرجات حرية $\nu = n-1$.

مثال (٧ – ١٩) إذا كان من المعتقد أن أكل السمك يساعد على زيادة الذكاء . أجريت تجربة على 11 شخصا تم اختيارهم عشوائيا وأجرى لهم أحد اختبارات الذكاء ثم أعطى لهم طعام يحتوى أساسا على السمك وبعد فترة معينة أجرى لهم اختبار الذكاء مرة أخرى فكانت النتائج كالتالى :

قبل أكل السمك	96	109	104	120	120	100	80	111	90	102	103
بعد أكل السمك	97	112	105	105	117	101	89	114	105	105	109

بفرض أن مستوى الذكاء قبل وبعد أكل السمك يتبع توزيعا طبيعيا وإذا اعتبرنا أن السمك له تأثير على مستوى ذكاء الأشخاص إذا لم تقع قيمة t المحسوبة من العينة في الفترة $(-t_{.05}\ ,\ t_{.05})$ بدرجات حرية $\nu = n-1 = 9$ ، ما الاستنتاج الذي يمكن الحصول عليه من بيانات العينة السابقة ؟

الحل .

من جدول t في ملحق (٤) فإن $t_{.05}= 1.812$ بدرجات حرية $\nu = 10$. وعلى ذلك يعتبر أكل السمك له تأثير على مستوى ذكاء الأشخاص إذا وقعت قيمة t خارج الفترة $(-1.812\ ,\ 1.812)$. من بيانات العينة فإن الفروق :

-1, -3, -1, 15, 3, -1, -9, -3, -15, -3, -6

ومنها :

$$n = 11, \Sigma d_i = -24, \Sigma d_i^2 = 606$$

$$\bar{d} = \frac{\Sigma d_i}{n} = \frac{-24}{11} = -2.1818,$$

$$s_{\bar{d}} = \sqrt{\frac{1}{n-1}\left[\Sigma d_i^2 - \frac{(\Sigma d_i)^2}{n}\right]}$$

$$= \sqrt{\frac{1}{10}\left[606 - \frac{(-24)^2}{11}\right]} = 7.44067.$$

$$t = \frac{\bar{d}}{s_d / \sqrt{n}} = \frac{2.1818}{7.44067 / \sqrt{11}} = 0.97252.$$

وحيث أن قيمة t المحسوبة تقع فى الفترة (1.812 , 1.812-) فيمكن القول أن السمك ليس له تأثير على مستوى الذكاء .

(٧-٧) توزيع مربع كاى **Chi - Square Distribution**

إذا تكرر سحب عينات من الحجم n من توزيع طبيعي تباينه σ^2 وإذا تم حساب تباين العينة s^2 لكل عينة فإننا نحصل على قيم للإحصاء S^2 . التوزيع العينى للإحصاء S^2 له تطبيقات قليلة في الإحصاء . اهتمامنا سوف يكون في توزيع المتغير X^2 والتي تحسب قيمته من الصيغة الآتية :

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \quad .$$

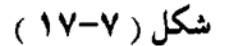
توزيع المتغير العشوائى X^2 يسمى توزيع χ^2 (توزيع مربع كاى) بدرجات حرية $v = n - 1$. كما ذكرنا سابقا فإن v تساوى المقام في صيغة s^2.

من الواضح أن قيم χ^2 لا يمكن أن تكون سالبة وعلى ذلك فإن منحنى توزيع χ^2 لا يمكن أن يكون متماثل حول الصفر. التوزيع العيني للإحصاء X^2 يمكن الحصول عليه باختيار عينات عشوائية متكررة من الحجم n من مجتمع طبيعي وحساب القيم χ^2 لكل عينه ثم يمكن الحصول على منحنى χ^2 بتمهيد المدرج التكرارى لقيم χ^2 . يعتمد شكل المنحنى على قيم v. يوضح شكل (٧-١٧) منحنيان لتوزيع χ^2 بدرجات حرية $v = 7$ و $v = 10$ حيث يمثل المنحنى بدرجات حرية $v = 7$ توزيع قيم χ^2 المحسوبة من كل العينات من الحجم $n = 8$ من مجتمع طبيعى تباينه σ^2 . بنفس الشكل يمثل المنحنى بدرجات حرية $v = 10$ توزيع كل قيم χ^2 المحسوب من كل العينات من الحجم $n=11$.

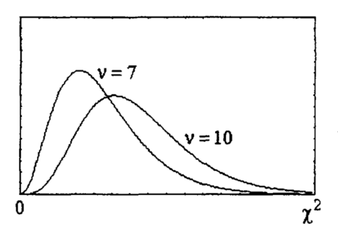
نظرية (٧-١٢) إذا كان s^2 يمثل تباين العينة من الحجم n المأخوذة من مجتمع طبيعى له تباين σ^2 فإن :

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

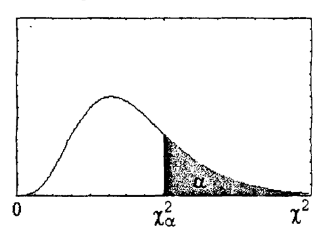
هي قيمة لمتغير عشوائى X^2 له توزيع χ^2 بدرجات حرية $v = n - 1$.



شكل (٧-١٧)



بفرض أن χ_{α}^{2} ترمز لقيمة χ^{2} التي توجد على المحور الأفقي تحت منحنى χ^{2} بدرجات حرية ν والتي تكون المساحة على يمينها قدرها α كما هو موضح في شكل (٧-١٨).



شكل (٧-١٨)

الجدول في ملحق (٥) يعطي قيم χ_{α}^{2} وذلك لقيم مختلفة من α و ν حيث α تأخذ القيم :

.995 , .99, .975, .95, .90, .10, .05, 0.025, .01, .005

ودرجات حرية من $\nu = 1$ إلى $\nu = 40$. يوضح الصف الثاني من الجدول قيم α والعمود الأول من الشمال قيم درجات الحرية أما محتويات الجدول فهي لقيم χ_{α}^{2}. وعلى ذلك للحصول على قيمة χ_{α}^{2} بدرجات حرية $\nu = 6$ والتي تكون المساحة على يمينها تساوى 05. فإننا نبحث في الجدول عند تقاطع الصف الذى به $\nu = 6$ مع العمود 05. وعلى ذلك $\chi_{.05}^{2} = 12.592$. ولعدم تماثل منحنى توزيع χ^{2} فلا بد من

استخدام الجدول لإيجاد $\chi^2_{.95} = 1.635$ عند $\nu = 6$.

مثال (٧ – ٢٠) أوجد قيمة χ^2 لتوزيع χ^2 بدرجات حرية $\nu = 14$ والتي تكون المساحة على يمينها تساوى **0.01** .

الحل .

بالبحث في جدول توزيع χ^2 في ملحق (٥) عند تقاطع الصف $\nu = 14$ مع العمود $\alpha = 0.01$ نجد أن $\chi^2 = 29.141$.

مثال (٧–٢١) أوجد قيمة χ^2 التي تكون المساحة على يسارها تساوى **0.99** لتوزيع χ^2 بدرجات حرية $\nu = 4$.

الحل .

قيمة χ^2 التي تكون المساحة على يسارها تساوي **.99** هي χ والتي تكون المساحة على يمينها تساوى **.01 = .99 -1** وعلى ذلك فإن قيمة $\chi^2_{.01}$ بدرجات حرية $\nu = 4$ هي تلك القيمة في جدول توزيع χ^2 التي تقع عند تقاطع الصف $\nu = 4$ والعمود $\alpha = .01$ هى $\chi^2_{.01} = 13.277$.

مثال (٧–٢٢) أوجد قيمتي χ^2 لتوزيع χ^2 بدرجات حرية $\nu = 15$ اللتين تحصران بينهما **99%** من المساحة تحت المنحنى بحيث أن المساحة في الطرف الأيسر للتوزيع تساوى المساحة في الطرف الأيمن للتوزيع.

الحل . المطلوب هنا هو إيجاد قيمتى χ^2 اللتين تقسمان المنحنى بحيث أن المساحة في الطرف الأيمن هى **.005**. وعلى ذلك قيمة χ^2 في الطرف الأيمن هى $\chi^2_{.005} = 32.799$. وللحصول على قيمة χ^2 في الطرف الأيسر والتي المساحة التي تقع على يسارها هى **0.005** وبالتالي فإن المساحة التى تقع على يمينها هى **.995 = .005-1** وعلى ذلك فإن قيمة χ^2 في الطرف الأيسر هى $\chi^2_{.995} = 4.600$ وبالتالي فإن القيمتين هما , **4.600** **32.799** .

(٧ –٨) توزيع F   F Distribution

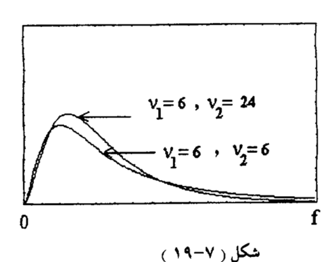
يعتبر توزيع **F** من التوزيعات الاحتمالية الهامة التى تستخدم في مجال الإحصاء التطبيقى . نظريا يمكن تعريف توزيع **F** (توزيع ف) كنسبة لتوزيعيين مستقلين يتبعان توزيع χ^2 وكل منهما له درجات حرية خاصة به . فـإذا كانت **f** قيمة للمتغير العشوائى **F** ، فإن:

وكل منهما له درجات حرية خاصة به . فإذا كانت f قيمة للمتغير العشوائى F ، فإن:

$$f = \frac{\chi_1^2 / \nu_1}{\chi_2^2 / \nu_2} = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} = \frac{\sigma_2^2 s_1^2}{\sigma_1^2 s_2^2}.$$

حيث χ_1^2 هي قيمة لتوزيع χ^2 بدرجات حرية $\nu_1 = n_1 - 1$ و χ_2^2 هي قيمة لتوزيع χ^2 بدرجات حرية $\nu_2 = n_2 - 1$.

لحساب قيمة f نختار عينة عشوائية من الحجم n_1 من مجتمع طبيعى له تباين σ_1^2 ونحسب s_1^2 / σ_1^2 . أيضا نختار عينة عشوائية مستقلة من الحجم n_2 من مجتمع طبيعى آخر له تباين σ_2^2 ونحسب s_2^2 / σ_2^2. النسبة للقيمتين s_1^2 / σ_1^2 و s_2^2 / σ_2^2 تنتج قيمة f. توزيع كل القيمة الممكنة من f حيث s_1^2 / σ_1^2 يمثل البسط و s_2^2 / σ_2^2 يمثل المقام يسمى توزيع F بدرجات حرية ν_1, ν_2. إذا اعتبرنا كل النسب الممكنة حيث s_2^2 / σ_2^2 يمثل البسط و s_1^2 / σ_1^2 يمثل المقام ، في هذه الحالة نحصل على توزيع كل القيم الممكنة التى تتبع توزيع F ولكن بدرجات حرية ν_2, ν_1. درجات الحرية المرتبطة بتباين القيمة التى في البسط دائما يوضع أولا متبوعا بدرجات الحرية المرتبطة بتباين العينة التى في المقام. وعلى ذلك منحنى توزيع F يعتمد ليس فقط على المعلمتين ν_1, ν_2. ولكن أيضا على ترتيبهما وبمجرد الحصول على القيمتين يمكن تعريف المنحنى. منحنيان لتوزيعF موضحة في شكل (٧-١٩).

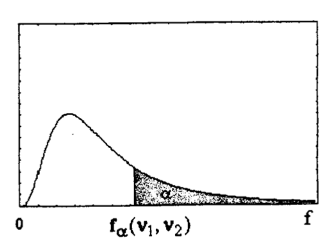


شكل (٧-١٩)

نظرية (٧-١٣) إذا كانت s_1^2 , s_2^2 تمثلان تبايني عينتين عشوائيتين مستقلتين من الحجم n_1 , n_2 مأخوذتين مجتمعين من طبيعيين بتباينى σ_1^2 , σ_2^2 على التوالي فإن :

$$f = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{\sigma_2^2 s_1^2}{\sigma_1^2 s_2^2}.$$

هي قيمة لمتغير عشوائي F يتبع توزيع F بدرجات حرية ν_1, ν_2 ، بفرض أن $f_\alpha(\nu_1, \nu_2)$ ترمز لقيمة f على المحور الأفقي تحت منحنى توزيع F بدرجات حرية $\nu_1 = n_1 - 1$ و $\nu_2 = n_2 - 1$ والتي تكون المساحة على يمينها تساوى α والموضحة في شكل (٧-٢٠).



شكل (٧-٢٠)

لاستخراج قيم $f_\alpha(\nu_1, \nu_2)$ يوجد جدولان في ملحق (٦) وملحق (٧) ، الأول عند $\alpha=$ .05 والآخر عند $\alpha = .01$ وفي كل منهما يكون الصف الأول لقيم ν_1 والعمود الأول لقيم ν_2 أما محتويات الجدول فهو لقيم $f_\alpha(\nu_1, \nu_2)$. على سبيل المثال من جدول توزيع F نلاحظ أن :

$$f_{.01}(5,7) = 7.46 \;,\; f_{.05}(1,4) = 7.71$$

$$f_{.01}(9,10) = 4.94 \;,\; f_{.05}(4,1) = 224.6$$

وباستخدام النظرية التالية يمكن استخدام جدول توزيع F في إيجاد $f_{1-\alpha}(\nu_1, \nu_2)$.

نظرية (٧-١٤) يمكن كتابة $f_{1-\alpha}(\nu_1, \nu_2)$ بدرجات حرية ν_1, ν_2 على الشكل :

$$f_{1-\alpha}(\nu_1,\nu_2)=\frac{1}{f_{\alpha}(\nu_2,\nu_1)}$$

وعلى ذلك فإن قيمة $f_{.95}(7,12)$ هى :

$$f_{.95}(7,12)=\frac{1}{f_{.05}(12,7)}=\frac{1}{3.57}=0.2801$$

حيث أن $f_{.05}(12\ ,\ 7)$ مستخرجة من جدول توزيع F في ملحق (٦) عند مستوى معنوية $\alpha = .05$ ودرجات حرية $\nu_1 = 12,\ \nu_2 = 7$ ،

تمارين :

-١- مجتمع محدود يتكون من القيم 2, 4, 6 (أ) أوجد المدرج التكراري للتوزيع العيني للإحصاء $\overline{X}$ عند سحب عينات عشوائية من الحجم n = 4 مع الإرجاع (ب) تحقق أن $\mu_{\overline{X}}=\mu$ و $\sigma^2_{\overline{X}}=\frac{\sigma^2}{n}$.

-٢- يتكون مجتمع من القيم 1, 1, 1, 2, 2, 3, 4 (أ) أوجد كل العينات الممكنة من الحجم n = 2 والتي يمكن اختيارها من هذا المجتمع بدون إرجاع (ب) أثبت أن :

$$\mu_{\overline{X}}=\mu \quad \text{و} \quad \sigma^2_{\overline{X}}=\frac{\sigma^2}{n}\,\frac{N-n}{N-1}\ .$$

-٣- إذا كانت 3, 4, 7, 9, 12 تمثل مفردات المجتمع محل الدراسة فإذا سحبت عينة مكونه من مفردتين :

(أ) أحسب التوقع والتباين للمجتمع .

(ب) حدد العينات الممكن سحبها بحيث يكون حجم كل عينة مفردتين إذا كان السحب مع الإرجاع .

(ج) حدد العينات الممكن سحبها إذا كان السحب بدون إرجاع .

(د) أوجد التوقع والتباين للعينات المسحوبة فى (ب) و (ج) .

-٤- كم عينة عشوائية مختلفة حجمها n = 3 يمكن اختيارها بإحلال ثم بدون إحلال من مجتمعات محدودة مكونة من N = 15, N = 20 , N = 25 ثم أوجد احتمال

اختيار كل عينة .

-٥- كم عينة عشوائية مختلفة حجمها $n = 2$ يمكن اختيارها بإحلال ثم بدون إحلال من مجتمعات محدودة مكونة من $N = 10$, $N = 15$, $N = 30$ ثم أوجد احتمال اختيار كل عينة .

-٦- ما هي قيمة معامل التصحيح للمجتمع المحدود عندما تكون :

(أ) $N = 100$, $n = 10$ (ب) $N = 200$, $n = 5$

(ج) $N = 50$, $n = 3$ (د) $N = 300$, $n = 15$

-٧- يقوم مصنع بإنتاج مصابيح كهربائية بمتوسط عمر 1900 ساعة وانحراف معياري 200 ساعة. احسب احتمال أن يكون متوسط العينة المبنى على عينة حجمها 100 مصباح أكبر من 1850 وأقل من 1920.

-٨- صممت آلة للشراب الرطب بحيث أن كمية الشراب الذى تخرجه يتبع التوزيع الطبيعى بمتوسط $\mu = 8$ أوقية للكوب وانحراف معياري 0.5 أوقية. تختبر الآلة دوريا بأخذ عينة من 9 أكواب وحساب متوسط العينة. إذا كان متوسط كمية الشراب المحسوب من 9 أكواب تقع في الفترة $\mu_{\overline{X}} \pm 2\sigma_{\overline{X}}$ فالآلة تعمل بطريقة صحيحة و غير ذلك لا بد من فحص الآلة واتخاذ اللازم . ما هو القرار الذى يجب اتخاذه عند سحب عينة عشوائية من 9 أكواب من الشراب الرطب وكان متوسط كمية الشراب للكوب 8.4 أوقية ؟

-٩- إذا كان ضغط الدم لمجموعة من الأفراد متوسطة 85 وانحرافه المعياري 5. أخذنا عينة عشوائية حجمها 40 فردا من هذه المجموعة، أوجد احتمال أن متوسط ضغط الدم في العينة أكبر من 80.

-١٠- إذا كان متوسط الدخل الشهري للأسر في إحدى المدن هو 7000 ريال بانحراف معياري 500 ريال. اختيرت عينة عشوائية حجمها 100 أسرة من هذه المدينة أوجد احتمال :

(أ) أن يقل متوسط دخل الأسرة في العينة عن 4800 ريال

(ب) أن يتراوح متوسط دخل الأسرة في العينة بين 6000 , 6500.

-١١- صممت إحدى الشركات سيارة بحيث أن أكبر حمولة لها 3000 كم ، وتتسع إلى 30 راكبا . إذا علمت أن أوزان الأشخاص الذين يستعملون هذه السيارة تخضع لتوزيع طبيعي متوسطة 70 كم وانحرافه المعياري 10 كم. أحسب احتمال أن تتحمل اكثر من طاقتها إذا كان مجموع الأشخاص الذين يستعملونها أكبر من 3000 كجم.

-١٢- إذا علمت أن درجات الطلاب في مقرر الإحصاء تخضع لتوزيع طبيعي وسطه 70 وانحرافه المعياري 9. أخذت عينة عشوائية حجمها 15 طالبا. أحسب احتمال أن يزيد متوسط علامات العينة عن 74.

-١٣- إذا كان X متغير عشوائيا وسطه 180.5 وانحرافه المعياري 12. من هذا التوزيع أخذت عينه حجمها 36. ما هو احتمال أن يزيد وسط العينة عن 190 ؟

-١٤- إذا كان X متغيرا عشوائيا يمثل أوزان أكياس الطحينة التى تنتجها إحدى المؤسسات وكان X خاضعا لتوزيع طبيعي وسطه 50 كيلو جراما وانحرافه المعيارى 5 كيلو جرام. أخذت عينة حجمها 9 أكياس من إنتاج هذه المؤسسة أحسب احتمال أن يزيد الوسط الحسابي لهذه العينة عن 51 كيلو جرام.

-١٥- إذا كان عدد الحوادث الأسبوعية على إحدى الطرق يتبع توزيع بواسون بمتوسط 5 حوادث فإذا أخذت عينة من 60 أسبوعا فما هو احتمال أن يكون متوسط عدد الحوادث فيها أقل من 3 حادث ؟

-١٦- إذا كان متوسط أعمار مرض السكر هو 60 سنة انحراف معياري 15 سنة. اختيرت عينة عشوائية حجمها 30 مريضا بالسكر. أوجد احتمال :

أ- أن يقل متوسط العمر في العينة عن 50 سنة .

ب - أن يتراوح متوسط العمر في العينة بين 40 ,50 سنة .

-١٧- سحبت عينة عشوائية من الحجم 36 من مجتمع كبير متوسطة $\mu = 80$ وانحرافه المعياري 18 فما هي النسبة المئوية من متوسطات العينات التي :

أ- تقع بين 71 و 89 . ب- أكبر من 89

ج- تقع بين 74 و 86 د- أقل من 74

-١٨- يرغب باحث في تقدير متوسط مجتمع وذلك بالاعتماد على عينة عشوائية مسحوبة من هذا المجتمع وتكون كبيرة بدرجة كافية وذلك تحت شرط :

$$P|\overline{X} - \mu| < \sigma / 4 = 0.95$$

ما هي أقل حجم للعينة يجب على الباحث استخدامه ؟

-١٩- إذا سحبت كل العينات الممكنة من الحجم 25 من مجتمع طبيعي بمتوسط 50 وانحراف معياري 5. ما هو الاحتمال أن متوسط العينة $\overline{X}$ سوف يقع في الفترة :

$$(\mu_{\overline{X}} - 1.96 \ \ \sigma_{\overline{X}}, \mu_{\overline{X}} + 1.96 \ \ \sigma_{\overline{X}})$$

-٢٠- إذا كانت أطوال 1000 طالب تقريبا تتبع التوزيع الطبيعي بمتوسط 68.5 بوصة وانحراف معياري 2.7 بوصة إذا سحبنا 200 عينة عشوائية من الحجم n =25 من هذا المجتمع وثم حساب المتوسطات . المطلوب :

أ- المتوسط والانحراف المعيارى للتوزيع العينى للمتوسط .

ب- عدد متوسطات العينات التى تقع في الفترة (68 , 69.2) .

ج- عدد متوسطات العينات التى تقل عن 67.2 .

-٢١- فتائل حريرية متوسط نقطة قطعها $\mu = 23$ رطلا بانحراف معيارى $\sigma = 0.4$ رطلا . اختيرت عينة عشوائية حجمها 40 فتيلة وذلك لإيجاد نقطة قطعها . ما هو احتمال أن متوسط نقطة القطع للعينة سيكون أكبر من 25.5 ؟

-٢٢- في دراسة عن تلوث الهواء بأكسيد الكبريت المنبعث من أحد المصانع كان متوسط التلوث $\mu = 18$ طنا بانحراف معياري $\sigma = 5.5$ طنا فإذا اختيرت عينة عشوائية مكونة من قراءات 90 يوما أحسب احتمال أن يكون متوسط العينة أكبر من 18.5 .

-٢٣- إذا كان X متغيرا عشوائيا يمثل أطوال الطلاب في جامعة ما ، وكان X خاضعا لتوزيع طبيعي متوسطة 160 سم وانحرافه المعيارى 20 سم . أخذت عينة عشوائية حجمها 15 طالبا من هذه الجامعة أحسب احتمال أن يقل متوسط العينة لأطوال أفراد هذه العينة عن 160 سم .

-٢٤- إذا كان $\overline{X}$ يمثل المتوسط لعينة من الحجم $n_1=2$ مع الإرجاع من مجتمع محدود له القيم 3, 4, 8 بنفس الشكل إذا كان $\overline{X}_2$ يمثل متوسط عينة من الحجم $n_2=2$ مع الإرجاع من مجتمع 2, 2, 4

أ- أوجد المدرج التكرارى للتوزيع العينى للإحصاء $\overline{X}_1 - \overline{X}_2$

ب - تحقق من أن $\mu_{\overline{X}_1-\overline{X}_2} = \mu_1 - \mu_2$, $\sigma^2_{\overline{X}_1-\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

-٢٥- سحبت عينة عشوائية من الحجم $n = 25$ من توزيع طبيعى لـه متوسـط 80 وانحراف معيارى 5. أيضا أخذت عينة عشوائية أخرى من الحجم 36 من مجتمع آخـر طبيعى له متوسط 75 وانحراف معيارى 3. أوجد الاحتمال أن متوسط العينة المحسوبة من 25 مفردة يزيد عن متوسط العينة المحسوب من 36 مفردة على الأقل 3.4 .

-٢٦- عينات عشوائية من الحجم 100 سحبت بدون إرجاع من مجتمعين A , B فإذا كان لدينا المعلومات التالية :

$$\mu_1 = 10 \ , \ \sigma_1 = 2 \ , \ \mu_2 = 8 \ , \ \sigma_2 = 1$$

أوجد أ- $E(\overline{X}_1 - \overline{X}_2)$ ب - $\sigma^2_{\overline{X}_1-\overline{X}_2}$

ج- الاحتمال أن الفرق $\overline{X}_1 - \overline{X}_2$ أقل من 1.5 .

د - الاحتمال أن الفرق $\overline{X}_1 - \overline{X}_2$ أكبر من 1.75 ولكن أقل من 2.5

-٢٧- أجرى بحث على أربعة طلبة لمعرفة هل يدخنون أم لا . فإذا كـانت نتـائج البحث كالآتى : No, yes, No, No

أ- أوجد نسبة المدخنين في المجتمع .

ب - أوجد التوزيع العينى للإحصاء $\hat{P}$ (نسبة المدخنـين) وذلـك في حالـة السحب بإرجاع أو بدون إرجاع ثم أوجد متوسط وتباين التوزيع إذا كان حجم العينـة التى تم اختيارها $n = 2$.

-٢٨- أجرى بحث على 5 أفراد لمعرفة حامل ميكروب معين فكانت النتائج كما يلي حامل للميكروب ، غير حامل ، حامل للميكروب ، غير حامل ، غير حامل المطلوب :

أ- نسبة الحاملين للميكروب في المجتمع .

ب - إذا اخترت نسبة عشوائية من فردين أوجد التوزيع العين للإحصاء $\hat{P}$ في

حالة السحب بإرجاع والسحب بدون إرجاع .

-٢٩- أجرى بحث لاستطلاع آراء 8 أفراد على منتج ما فكانت النتائج :

yes , yes , yes , yes , yes , No, No, No المطلوب :

أ- نسبة الأفراد المفضلون للمنتج .

ب - إذا اختيرت عينة عشوائية من فردين أوجد التوزيع العينى للإحصاء $\hat{P}$ في حالة السحب بإرجاع والسحب بدون إرجاع .

-٣٠- إذا كانت نسبة المدخنين من الأفراد الذكور البالغين في إحدى المدن 20% فإذا اختيرت عينة عشوائية من 200 شخص أوجد احتمال أن يقل عدد المدخنين بينهم عن 3 أشخاص .

-٣١- في شركة كبيرة من 4000 موظف وجد أن الغياب في يوم الاثنين يمثل 8% لعدة سنوات. إذا اختيرت عينة عشوائية من 1500 موظف في يوم الاثنين. أوجد احتمال أن تكون نسبة الغياب أكبر من 0.07 .

-٣٢- إذا كان 45% من الأفراد من أصحاب الأملاك في بلد كبير يملكون سيارتين على الأقل. ما هو الاحتمال في عينة عشوائية من 150 شخص من هذا البلد تكون نسبة امتلاك سيارتين على الأقل 0.55 .

-٣٣- في مجتمع معين كانت نسبة المعاناة من مرض ما هو 0.2. أخذت عينة عشوائية من هذا المجتمع حجمها 250 (السحب بإرجاع) أوجد :

أ- احتمال أن تكون النسبة في هذه العينة أقل من 0.3 .

ب - بفرض أن السحب بدون إرجاع وحجم المجتمع 500 أوجد احتمال أن تكون النسبة في العينة أكبر من 0.4 .

-٣٤- في دولة ما معروف أن 70% من المرشحين في الانتخابات يناهضون قضية ما. فإذا اختيرت عينة عشوائية من هذه الدولة مكون من 100 مرشح أوجد احتمال أن تكون نسبة المرشحين الذين يناهضون القضية في العينة بين 80% , 60% .

-٣٥- إذا علم أن نسبة البيض التالف التى تنتجها أحد مراكز إنتاج البيض هى 0.2 . أشترى شخص 300 بيضة من إنتاج هذا المركز أوجد احتمال أن يجد من بينها 10 بيضة على الأقل تالفة .

-٣٦- إذا كان لدينا آلتين وكانت نسبة المعيب للآلة الأولى هو 15% والمعيب للآلة الثانية هو 12 % . سحبت عينتان من إنتاج الآلتين حجمهما 650 ,600 وحدة على التوالي . فإذا كانت نسبة المعيب للعينة الأولى هو $\hat{P}_1$ ونسبة المعيب في العينة الثانية $\hat{P}_2$. احسب قيمة الاحتمال $P(\hat{P}_1 - \hat{P}_2 \leq 0.2)$.

-٣٧- أوجد : أ - $t_{0.99}$, $\nu = 10$ ب - $t_{.05}$, $\nu = 15$

ج- $t_{.025}$, $\nu = 17$ د - $t_{.99}$, $\nu = 18$

-٣٨- أوجد قيمة t_α عندما $\nu = 23$ بحيث أن $P(-t_\alpha < T < t_\alpha) = .99$.

-٣٩- أوجد قيمة α التي تحقق الآتي : $t_\alpha = 1.33$ و $\nu = 18$.

-٤٠- أوجد الاحتمال أن المتغير العشوائي T والذى يتبع توزيع t :

أ- أكبر من 1.74 عند $\nu = 17$ ب - أقل من 1.323 عند $\nu = 21$

-٤١- يقوم مصنع بإنتاج مصابيح كهربائية متوسط عمرها 500 ساعة . للمحافظة على هذا الإنتاج يقوم المصنع شهريا باختيار 25 مصباحا. إذا وقعت قيمة t المحسوبة في الفترة $(-t_{.05} , t_{.05})$ فإن الإنتاج يكون محققا للمواصفات القياسية. ما هو الاستنتاج الذى يمكن اتخاذه عند اختيار عينة لها متوسط $\bar{x} = 518$ وانحراف معياري $s=40$ ساعة ؟ وذلك تحت فرض أن العينة تم اختيارها من مجتمع طبيعي .

-٤٢- أوجد المئين العاشر والمئين التسعين P_{90} لتوزيع t بدرجات حرية $\nu = 15$.

-٤٣- أوجد الربيع الأول Q_1 والربيع الثالث Q_3 والوسيط Q_2 لتوزيع t بدرجات حرية $\nu = 20$.

-٤٤- إذا كانت الأجور اليومية لعمال إحدى الشركات في مدينة ما تخضع للتوزيع الطبيعي بمتوسط $\mu_1 = 5$ وانحراف معياري $\sigma_1 = 0.5$. وكانت الأجور اليومية لعمال شركة مماثلة في مدينة أخرى تخضع للتوزيع الطبيعي بمتوسط $\mu_2 = 4$ وانحراف معياري $\sigma_1 = 10$ وبفرض أننا سحبنا عينة عشوائية من المجتمع الأول من الحجم $n_1 = 25$ وعينة أخرى من المجتمع الثاني من الحجم $n_2 = 20$ أوجد $P(\bar{X}_1 - \bar{X}_2 \leq 0.75)$.

-٤٥- اختيرت عينة عشوائية من مجتمع طبيعي له متوسط $\overline{x}_1 = 20$ وانحراف معياري $s_1 = 2$. كما اختيرت عينة أخرى عشوائية من الحجم $n_2 = 12$ من مجتمع آخر له متوسط $\overline{x}_2 = 24$ وانحراف معياري $s_2 = 6$ فإذا كانت $\mu_2 = 19$, $\mu_1 = 22$ و σ_1^2 , σ_2^2 مجهولتان ولكن تقريبا متساويتان أوجد قيمة t .

-٤٦- طبق اختبار للقدرة على التفكير الناقد على مجموعة من المراهقين قبل حضورهم برنامج أعد لهذا الفرض مدته 40 أسبوعا وبعد حضورهم للبرنامج. فإذا اختيرت عينة عشوائية من الحجم $n = 10$ وكان متوسط الفروق $\overline{d} = 5$ بانحراف معياري $s_{\overline{d}} = 0.5$ أوجد قيمة t .

-٤٧- أوجد النقاط التالية من جدول توزيع χ^2 في ملحق (٥) مع التوضيح بالرسم :

أ- $\chi^2_{.95}$, $\nu = 12$ ب- $\chi^2_{.05}$ و $\nu = 1$

-٤٨- لتوزيع χ^2 أوجد

أ- $\chi^2_{.001}$, $\nu = 18,$ ب- $\chi^2_{.975}$, $\nu = 29,$

ج- χ^2_{α} بحيث أن $P(X^2 \le \chi^2_{\alpha}) = .99$, $\nu = 4$.

-٤٩- أوجد النقاط التالية من جدول توزيع F في ملحق (٦) مع التوضيح بالرسم $f_{.95}(6, 10)$, $f_{.05}(12,7)$.

-٥٠- أوجد الاحتمال أن المتغير العشوائي F أكبر من 2.85 إذا كانت F له توزيع F بدرجات حرية $\nu_1 = 8, \nu_2 = 12$.

-٥١- أوجد القيمة $f_\alpha(12,8)$ إذا كانت $P(F < f_\alpha(12,8)) = 0.05$.

-٥٢- أوجد القيمة α لكل من القيم التالية مع التوضيح بالرسم $f_\alpha(3,20) = 3.1, \quad f_\alpha(9,8) = 3.39$.

-٥٣- أوجد الاحتمال أنه من عينة عشوائية من 25 مفردة من مجتمع طبيعي له تباين $\sigma^2 = 6$ سوف يكون التباين S^2

أ - أكبر من 9.1 ب- بين 4.62 , 10.745

٥٤- إذا كانت s_2^2, s_1^2 تمثلان تباين عينتين عشوائيتين من الحجم $n_1 = 31$, $n_2 = 25$ تم اختيارهما من مجتمعين طبيعيين . بحيث $\sigma_1^2 = 10$, $\sigma_2^2 = 15$ أوجد :
$P(S_1^2 / S_2^2) > 1.26)$.

٥٥- إذا كانت s_2^2, s_1^2 تمثلان تباين عينتين عشوائيتين من الحجم $n_1 = 8$, $n_2 = 12$ مسحوبتين من مجتمعين طبيعيين تحت فرض أن $\sigma_1^2 = \sigma_2^2$. أوجد الاحتمال
$P(S_1^2 / S_2^2 < 4.89)$.

الفصل الثامن

فترات الثقة

Confidence Intervals



(٨-١) مقدمـــة Introduction

يعتبر الاستدلال الإحصائي statistical inference فرع في علم الإحصـــاء يـــهتم بطرق الاستدلال أو التعميم بشان المجتمع وذلك بالاعتماد على معلومات يتم الحصول عليها من عينات مختارة من المجتمع . سوف نتناول في هذا الفصل الاستدلال عن معالم مجتمعات مجهولة مثل المتوسط ، النسبة ، الانحراف المعياري ، وذلك بحساب إحصاءات من عينات عشوائية وتطبيـــق نظرية المعاينة التي ناقشناها في الفصل السابع .

ينقسم فرع الاستدلال الإحصائي إلى فرعين أساسين : التقدير estimation واختبارات الفروض tests of hypotheses . سوف تقتصر دراستنا في هذا الفصـــل علـــى موضـــوع التقدير بينما موضوع اختبارات الفروض سوف نتناوله في الفصل التاسع . الأمثلة التالية توضح الفرق بين الفرعين. يقوم مصنع بإنتاج قضبانا حديدية ، فإذا اختيرت عينة عشوائية مكونه مـــن 200 قضيب من إنتاج هذا المصنع وقيست أطوالها وتم حساب متوسط طول القضيب في العينة . هذا المتوسط يمكن أن يستخدم لتقدير المعلمة الحقيقية للمجتمع μ . المعلومات عـــن توزيـــع المعاينة للإحصاء $\overline{X}$ سوف يساعدنا في حساب درجة الثقة في تقديرنا. هذه المشكلة تنتمـــي إلى فرع التقدير . الآن إذا كان معروفا أن جسم الإنسان البالغ يحتاج يوميا في المتوسـط إلى 800 ملليجرامات من الكالسيوم لكي يقوم بوظائفه خير قيام. يعتقد علماء التغذية أن الأفـــراد ذوى الدخل المنخفض لا يستطيعون تحقيق هذا المتوسط . لاختبار ذلك اختيرت عينة عشوائية من 50 شخصا بالغا من بين ذوى الدخل المنخفض وتم حساب متوسط ما يتناولونه من الكالسيوم يوميا. في هذا المثال لم نحاول تقدير معلمة ولكن بدلا من ذلك نحاول الوصول إلى قرار صحيـــح عـــن الفرض الذي وضعه علماء التغذية . مرة أخرى نعتمد على نظرية المعاينة لتمدنا بمقياس للدرجـــة الثقة في القرار الذي نتخذه.

يتم تقدير معلمة المجتمع إما كتقدير بنقطة point estimate أو كتقديـــر بفـــترة interval estimate . تقدير النقطة لمعلمة مجتمع ما θ هي قيمـــة وحيـــدة (مفـــردة) $\hat{\theta}$ للإحصاء $\hat{\Theta}$. على سبيل المثال القيمة $\bar{x}$ للإحصاء $\overline{X}$ ، والمحسوبة من عينة عشـــوائية مـــن الحجم n ، هي تقدير بنقطة لمعلمة المجتمع μ . بنفس الشكل ، $\hat{p} = \frac{x}{n}$ هي تقديـــر بنقطـــة للمعلمة الحقيقية p والتي تمثل نسبة صفة ما في مجتمع .

الإحصاء المستخدم لإيجاد تقدير النقطة يسمى المقـــدر estimator أو دالـــة القـــرار decision function. فعلى سبيل المثال دالة القرار S ، والتي تكـــون دالـــة في العينـــة العشوائية ، هي مقدر للمعلمة σ . عينات مختلفة تؤدى إلى تقديرات مختلفة .

بفرض أن $\hat{\Theta}$ مقدر حيث القيمة $\hat{\theta}$ هي تقدير بنقطة لمعلمة مجتمع مجهولة θ. من المؤكد أننا نرغب في إيجاد التوزيع العيني للإحصاء $\hat{\Theta}$ والذي متوسطة يساوى المعلمة التي نرغـــب في تقديرها. أي مقدر يحقق هذه الخاصية يسمي مقدر غير متحيز unbiased estimator .

تعريف : يقال للإحصاء $\hat{\Theta}$ أنه مقدر غير متحيز للمعلمة θ إذا كان $E(\hat{\Theta})=\theta$. إذا كان $\hat{\Theta}_1, \hat{\Theta}_2$ مقدران غير متحيزان لمعلمة مجتمع θ فإننا نختار المقدر الذي توزيعه الاحتمالى له أقـــل تباين . وعلى ذلك إذا كان $\sigma^2_{\hat{\Theta}_1} < \sigma^2_{\hat{\Theta}_2}$ ، فإننا نقول أن $\hat{\Theta}_1$ مقدرا أكثر كفاءة من $\hat{\Theta}_2$.

تعريف : اعتبر كل المقدرات الغير متحيزة لمعلمة θ . يسمي المقدر الذي له أقل تباين بـــالمقدر الأكثر كفاءة more efficient للمعلمة θ . للتوزيع الطبيعي تم إثبات أن كلا من $\bar{X}, \tilde{X}$ (الوسيط) مقدران غير متحيزين لمعلمة المجتمع μ ولكن تباين $\bar{X}$ أقل من تباين $\tilde{X}$. وعلـــى ذلك كلا التقديرين $\bar{x}, \tilde{x}$ (وسيط العينة) ، في المتوسط ، سوف يكون لهما نفس متوسط المجتمع μ ، ولكن $\bar{x}$ سوف يكون أقرب للمعلمة μ من $\tilde{x}$.

أي تقدير بفترة لمعلمة θ هو فترة على الشكل $a < \theta < b$ حيث a , b تعتمدان على التقدير بنقطة $\hat{\theta}$ لعينة عشوائية خاصة مختارة من المجتمع موضع الدراسة وأيضا علـــى التوزيـــع العيني للإحصاء $\hat{\Theta}$. على سبيل المثال إذا اختيرت عينة عشوائية تمثل درجـــات التحصيـــل في امتحان القبول لخمسين طالبا من المتقدمين للالتحاق في كلية ما وتم الحصول على الفترة , 500 550 والتي نتوقع أن المتوسط الحقيقي لدرجات التحصيل داخلها . القيمتان النهائيتان 500 و 550 سوف تعتمدان على متوسط العينة المحسوبة $\bar{x}$ وأيضا على التوزيع العيـــني $\bar{X}$. كلمـــا زادت حجم العينة ، فإن $\sigma^2_{\bar{X}} = \dfrac{\sigma^2}{n}$ سوف تقل ، كما عرفنا في الفصل السابع ، وبالتالى فـــإن تقديرنا سوف يقترب من المعلمة μ ويؤدى إلى فترة قصيرة .

عينات مختلفة تؤدى إلى قيم مختلفة لـــ $\hat{\theta}$ وبالتالى إلى تقديرات بفترة لمعلمة المجتمـــع θ . بعض هذه الفترات سوف تحتوى على θ والبعض الآخر لا يحتوى علـــى θ. التوزيـــع العيـــني للإحصاء $\hat{\Theta}$ سوف يساعدنا في إيجاد a , b لكل العينات الممكنة بحيث أن أي نسبة خاصة مـــن هذه الفترات سوف تحتوى على θ . فعلى سبيل المثال ، يتم حساب a , b بحيث تكون 0. 95 من كل الفترات الممكنة ، مع تكرار المعاينة ، سوف تحتوي على θ . وعلى ذلك يكون لدينـــا

احتمال 0.95 لاختيار واحدة من هذه العينات والتي تؤدى إلى فترة تحتوي على θ. هذه الفــترة المحسوبة من عينة عشوائية ، تسمي %95 فترة ثقة confidence interval . بمعنى آخـــر يكون لدينا %95 ثقة أن فترتنا المحسوبة تحتوى على المعلمة θ. عموما توزيع $\hat{\Theta}$ ســوف يساعدنا في حساب a , b بحيث يكون لأي نسبه خاصة $1-\alpha$, $0<\alpha<1$ ، من الفــترات المحسوبة من كل العينات الممكنة سوف تحتوى علــى المعلمــة θ . الفــترة المحســوبة تســمى $(1-\alpha)100\%$ فترة ثقة للمعلمة θ . تعتبر فترة الثقة الأطول ، هي الأكثر ثقة في الحصـــول على فترة تحتوي على المعلمة المجهولة . بالطبع يكون من الأفضل الحصول على %95 فترة ثقــة أن متوسط العمر لنوع معين من البطاريات ينحصر بين 8 و 5 أسابيع عن الحصول على %99 فترة ثقة أن متوسط العمر ينحصر بين 11 و 2 أسبوعا . دائما يفضل الحصول على فترة قصيرة بدرجة عالية من الثقة . في البنود التالية من هذا الفصل سوف نتناول المقدرات الأكثر كفـــاءة لمعالم المجتمع الشائعة الاستخدام. باستخدام هذه المقدرات وتوزيعاتها الاحتمالية ، يمكن اشـــتقاق فترات الثقة المقابلة لها.

(٨-٢) فترة ثقة لمتوسط المجتمع μ

Confidence Interval for Population Mean μ

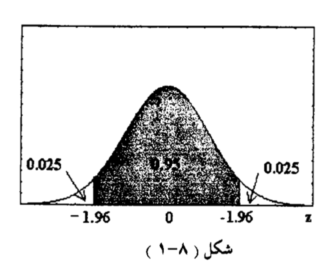
يعتبر الإحصاء $\bar{X}$ هو المقدر الأكثر كفاءة لمعلمة المجتمع μ . التوزيع العيني للإحصـــاء $\bar{X}$ يرتكز عند μ وتباينه أقل من أي مقدر آخر. وعلى ذلك فمتوســـط العينـــة $\bar{x}$ ســوف يستخدم كتقدير بنقطة لمتوسط المجتمع μ . مرة أخرى فإن $\sigma^2_{\bar{X}}=\dfrac{\sigma^2}{n}$ وعلى ذلـــك فـــإن العينة الكبيرة سوف تؤدى إلى قيمة للإحصاء $\bar{X}$ ناتجة من توزيع عيني له تباين صغـــير. وعلـــى ذلك فإن $\bar{x}$ تقترب من μ عندما تكون n كبيرة .

الآن سوف نتناول الخطوات المتبقية في إيجاد فترة ثقة لمتوسط المجتمع μ وذلـــك تحـــت فرض أن العينة مختارة من مجتمع طبيعي أو، عند عدم تحقق هذا الفرض ،إذا كانت n كبـــيرة بدرجة كافية. وتبعا للنظريتين (٧-٣) و (٧-٥) فإننا نتوقع أن التوزيع العيني للإحصـــاء $\bar{X}$ تقريبا يتبع التوزيع الطبيعي بمتوسط $\mu_{\bar{X}}=\mu$ وتباين $\sigma^2_{\bar{X}}=\dfrac{\sigma^2}{n}$. تبعا لذلك فإن المتغـــير العشوائي الطبيعي القياسي Z سوف ينحصر بين 1.96- , 1.96 وذلك باحتمـــال 0.95 حيث :

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}.$$

يتضح من شكل (٨-١) أن :

$$P(-1.96 < Z < 1.96) = 0.95.$$



شكل (٨-١)

وبالتعويض عن Z بقيمتها وبضرب كل حد في المتباينة في $\frac{\sigma}{\sqrt{n}}$ وطرح $\overline{X}$ من كل حد والضرب في 1- نحصل على :

$$P(\overline{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + 1.96\frac{\sigma}{\sqrt{n}}) = 0.95 .$$

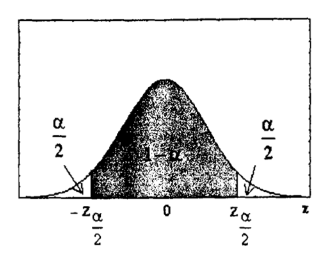
وعلى ذلك يمكن القول أنه باحتمال 95% تحتوى الفترة العشوائية $\overline{X} \pm \frac{1.96\sigma}{\sqrt{n}}$ على المعلمة μ .وعلى ذلك نختار عينة عشوائية من الحجم n من المجتمع الذي تباينه σ^2 معلوم ونحسب متوسط العينة $\overline{x}$ وذلك للحصول على 95% فترة ثقة على الشكل :

$$\overline{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + 1.96\frac{\sigma}{\sqrt{n}} .$$

حيث 1.96 تمثل القيمة الحرجة المقابلة لمعامل الثقة 0.95 وحدي الثقة هما $\overline{x} + 1.96\sigma/\sqrt{n}$, $\overline{x} - 1.96\sigma/\sqrt{n}$ حيث σ معلومة.

الآن سوف نشرح بصورة عامة طريقة الحصول على $(1-\alpha)100\%$ فترة ثقة حيث $0<\alpha<1$. يتضح من شكل (٨-٢) أن $z_{\frac{\alpha}{2}}$ تمثل قيمة z التي تكون المساحة على يمينها تساوي $\frac{\alpha}{2}$. وعلى ذلك يمكن كتابة :

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1-\alpha \ .$$



شكل (٨-٢)

وبالتعويض عن Z بقيمتها وبضرب كل حد في المتباينة في $\frac{\sigma}{\sqrt{n}}$ وطرح $\overline{X}$ من كل حد والضرب في 1- نحصل على :

$$P(\overline{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}) = 1-\alpha \ .$$

وعلى ذلك نختار عينة عشوائية من الحجم n من المجتمع الذي تباينه σ^2 معلوم ونحسب متوسط العينة $\overline{x}$ وذلك للحصول على $(1-\alpha)100\%$ فترة ثقة على الشكل :

$$\overline{x} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \ .$$

للعينات العشوائية الصغيرة المختارة من مجتمعات غير طبيعية ، لا نتوقع أن درجة ثقتنا تكون مضبوطة. للعينات من الحجم $n \geq 30$ وبصرف النظر عن شكل المجتمع فإن نظرية المعاينة تؤمن لنا نتائج جيدة.

لحساب $(1-\alpha)100\%$ فترة ثقة للمعلمة μ نفترض أن σ معلومة ولكن عموما لا يتوافر هذا الفرض ، في هذه الحالة يمكن الاستعاضة عن σ بالانحراف المعياري للعينة s بشرط أن $n \geq 30$.

مثال (٨-١) اختيرت عينة عشوائية من 100 سيجارة من نوع معين وكان متوسط النيكوتين فيها 25 ملليجراما والانحراف المعياري هو 6 ملليجرامات . أوجد فترة ثقة لمتوسط النيكوتين في السجائر من هذا النوع بدرجة ثقة 95% و 99% .

الحل . التقدير بنقطة للمعلمة μ هو $\bar{x} = 25$ وحيث أن حجم العينة كبير ، فإن الانحراف المعياري للمجتمع σ يمكن الاستعاضة عنه بالانحراف المعياري للعينة $s=6$. من جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن قيمة z التي على يمينها مساحة قدرها 0.025 وعلى يسارها مساحة قدرها 0.975 هي $z_{0.025} = 1.96$. وعلى ذلك فإن 95% فترة ثقة سوف تكون على الشكل :

$$25 - \frac{(1.96)(6)}{\sqrt{100}} < \mu < 25 + \frac{(1.96)(6)}{\sqrt{100}}$$

والتي تختزل إلى :

$$23.824 < \mu < 26.176$$

للحصول على 99% فترة ثقة وباستخدام جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن قيمة z التي على يمينها مساحة قدرها 0.005 وعلى يسارها مساحة قدرها 0.995 هي $z_{.005}=2.575$ وعلى ذلك فإن 99% فترة ثقة تكون على الشكل :

$$25 - \frac{(2.575)(6)}{\sqrt{100}} < \mu < 25 + \frac{(2.575)(6)}{\sqrt{100}}$$

والتي تختزل إلى :

$$23.455 < \mu < 26.545 .$$

تمدنا $(1-\alpha)100\%$ فترة الثقة بتقدير لدقة التقدير بنقطة الذي نحصل عليه. إذا وقعت μ عند مركز الفترة فإن $\bar{x}$ تقدر μ بدون أخطاء. في معظم الحالات فإن $\bar{x}$ لا تساوى μ وبالتالي نحصل على تقدير بنقطة ، $\bar{x}$ ، بخطأ error . حجم هذا الخطأ يساوي الفرق بين $\bar{x}$ و μ. يصل هذا الخطأ إلى اقصاة عندما تكون μ قريبة من إحدى حدي الثقة. أي أن $\bar{X}$ سوف تختلف عن μ بمقدار أقل من $z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ كما يتضح من شكل (٨-٣) .

Error

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \qquad \bar{x} \quad \mu \qquad \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

شكل (٨-٣)

نظرية (٨-١) إذا استخدمت $\bar{x}$ كتقدير لمعلمة المجتمع μ فإنه يكون لدينـــا $(1-\alpha)100\%$ ثقة أن الخطأ سوف يكون أقل من $z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$.

في المثال (٨-١) يكون لدينا 99% ثقة أن متوسط العينة $\bar{x} = 25$ يختلف عـــن المتوســـط الحقيقي μ بمقدار أقل من 1.545 .

عادة نرغب في معرفة حجم العينة اللازم للتأكد من أن الخطأ في تقدير μ سوف يكون اقل من قيمة e . تعنى نظرية (٨-١) أننا لا بد من اختيار n بحيث $e = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$.

نظرية (٨-٢) إذا استخدمت $\bar{x}$ كتقدير للمعلمة μ فإن يكون لدينا $(1-\alpha)100\%$ ثقة أن الخطأ سوف يكون أقل من قيمة معينة e وذلك عندما يحسب حجم العينة من الصيغة التالية :

$$n = \left(\frac{z_{\frac{\alpha}{2}} \sigma}{e} \right)^2.$$

الصيغة السابقة تمكن المرء في تحديد مدي كبر العينة التي يحتاج إليها كي يقدر μ لأي درجـــة يرغبها من درجات الدقة قبل أخذ أي عينة واحدة شريطة أن تكون قيمة σ معلومة . أمـــا إذا لم يكن المرء على علم بقيمة σ ، فلا بد من أخذ عينة مبدئية ، $n \geq 30$ ،كي يحصل على تقديـــر للمعلمة σ يمكن استخدامه في الصيغة السابقة لتحديد مدي كبر n الواجب .

مثال (٨-٢) ما هو حجم العينة المطلوب في مثال (٨-١) وذلك للحصول على 95% ثقة أن تقديرنا الذي نحصل عليه يختلف عن μ بقيمة أقل من واحد صحيح.

الحل . الانحراف المعياري $s = 6$ والمحسوب من عينة عشوائية من الحجم $n = 100$. سوف تستخدم s بدلا من σ . من نظرية (٨-٢) فإن :

$$n = \left[\frac{(1.96)(6)}{1}\right]^2 = 138.3 \approx 138.$$

وعلى ذلك يمكن القول أن لدينا 95% ثقة أن العينة العشوائية من الحجم 138 سوف تمدنا بتقدير $\bar{x}$ يختلف عن μ بقيمة أقل من الواحد الصحيح .

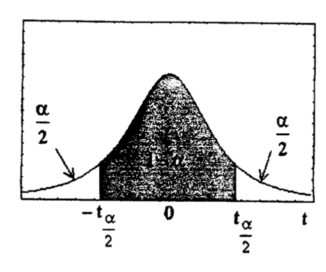
في معظم الأحيان يكون المطلوب تقدير متوسط المجتمع عندما يكون التباين غـــير معلـــوم وحجم العينة أقل من 30، فقد تكون التكاليف عاملا محددا لحجم العينة. طالما كان شكل المجتمع (تقريبا) ناقوسى فإنه يمكن حساب فترات الثقة عندما تكون σ^2 غير معلومة وحجم العينة صغير وذلك باستخدام التوزيع العيني للمتغير T ، حيث أن :

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}.$$

طريقة إيجاد $(1-\alpha)100\%$ فترة ثقة في هذه الحالة هي نفسها الطريقة المتبعة في حالة العينـــات الكبيرة فيما عدا استخدام توزيع t بدلا من التوزيع الطبيعي القياسي.

يتضح من شكل (٨-٤) أن :

$$P(-t_{\frac{\alpha}{2}} < T < t_{\frac{\alpha}{2}}) = 1 - \alpha.$$



شكل (٨-٤)

حيث أن $t_{\frac{\alpha}{2}}$ هي قيمة t بدرجة حرية $v = (n-1)$ والتي تكون المساحة على يمينها تساوي $\frac{\alpha}{2}$.

ونظرا لخاصية التماثل لتوزيع t فإن مساحة مساوية قدرها $\frac{\alpha}{2}$ تقع على يسار القيمة $-t_{\frac{\alpha}{2}}$.

بالتعويض عن T بقيمتها فإننا يمكن كتابة :

$$P(-t_{\frac{\alpha}{2}} < \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} < t_{\frac{\alpha}{2}}) = 1 - \alpha.$$

وبضرب كل حد في المتباينة في $\frac{S}{\sqrt{n}}$ وطرح $\overline{X}$ من كل حد والضرب في 1- نحصل على :

$$P(\overline{X} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}) = 1 - \alpha \ .$$

لعينة خاصة من الحجم n ، يحسب المتوسط $\overline{x}$ والانحراف المعياري s ويتم الحصول على $(1-\alpha)100\%$ فترة ثقة كما يأتي :

$$\overline{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \overline{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \ .$$

مثال (٨-٣) في اختبار للزمن الذي يستغرقه تجميع ماكينة معينة وجد أن الزمن الذي استغرقه تجميع 6 ماكينات هو على التوالي : 12, 13, 11, 5, 10, 12 (مقاسه بالدقائق). أوجد 95% فترة ثقة لمتوسط المجتمع μ وذلك تحت فرض أن الزمن (في هذا المثال) يتبع توزيعاً طبيعياً.

الحل . متوسط العينة والانحراف المعياري للبيانات المعطاة هما : $s = 2.881$, $\overline{x} = 10.5$.

باستخدام جدول توزيع t في ملحق (٤) فإن $t_{0.025} = 2.571$ وذلك عند درجات حريه = v 5 . وعلى ذلك 95% فترة ثقة للمعلمة μ هي :

$$10.5 - \frac{(2.571)(2.881)}{\sqrt{6}} \ < \ \mu \ < \ 10.5 + \frac{(2.571)(2.881)}{\sqrt{6}}$$

التي تختزل إلى :

$$7.476 \ < \ \mu \ < \ 13.524 .$$

(٨-٣) فترة ثقة للفرق بين متوسطي مجتمعين $\mu_1 - \mu_2$

Confidence Interval for the Difference Between two Populations Means

إذا كان لدينا مجتمعان ، المجتمع الأول له متوسط μ_1 وتباين σ_1^2 والمجتمع الثـــاني لـــه متوسط μ_2 وتباين σ_2^2. التقدير الأكثر كفاءة للفرق بين متوسطين $\mu_1 - \mu_2$ هو الإحصاء $\overline{X}_1 - \overline{X}_2$. وعلى ذلك ، للحصول على تقدير بنقطة للمعلمة $\mu_1 - \mu_2$ لا بد من اختيـــار عينة عشوائية من الحجم n_1 من المجتمع الأول وعينة عشوائية من الحجم n_2 من المجتمـــع الثـــاني ومستقلة عن العينة الأولى وحساب الفرق بين المتوســـطين $\overline{x}_1 - \overline{x}_2$. بفـــرض أن العينتـــين المستقلين تم اختيارهما من مجتمعين طبيعيين ، أو في حالة عدم توافر ذلك الفرض ، إذا كان كـــلا من n_1 و n_2 أكبر من أو يساوي 30 فإنه يمكن إيجاد فترة ثقة للمعلمة $\mu_1 - \mu_2$ تعتمد علـــى التوزيع العيني للإحصاء $\overline{X}_1 - \overline{X}_2$.

بالرجوع إلى نظرية (٧-٦) فإننا نتوقع أن التوزيع العيني للإحصاء $\overline{X}_1 - \overline{X}_2$ تقريبا يتبـــع التوزيع الطبيعي بمتوسط :

$$\mu_{\overline{X}_1-\overline{X}_2} = \mu_1 - \mu_2$$

وانحراف معياري :

$$\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

وعلى ذلك فإن المتغير العشوائي الطبيعي القياسي :

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

سوف يقع بين $z_{\frac{\alpha}{2}}$ و $-z_{\frac{\alpha}{2}}$ باحتمال $1-\alpha$. وبالرجوع مرة أخـــرى إلى شـــكل (٨-٢) فإنه يمكن كتابة :

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1 - \alpha .$$

وباستبدال Z بقيمتها فإن :

$$P\left[-z_{\frac{\alpha}{2}} < \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} < z_{\frac{\alpha}{2}}\right]=1-\alpha.$$

وبضرب كل حد في المتباينة في $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ وطرح $\overline{X}_1-\overline{X}_2$ من كل حد والضرب في 1 - نحصل على :

$$P\left[(\overline{X}_1-\overline{X}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} < \mu_1-\mu_2 <\right.$$

$$\left.(\overline{X}_1-\overline{X}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]=1-\alpha.$$

لأى عينتين عشوائيتين مستقلتين من الحجم n_2 , n_1 مأخوذتين من مجتمعين تباينهما σ_1^2 , σ_2^2 معلومتين ، فإن $\overline{x}_1-\overline{x}_2$ تحسب ويمكن إيجاد $(1-\alpha)100\%$ فترة ثقة كالآتي :

$$(\overline{x}_1-\overline{x}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} < \mu_1-\mu_2$$

$$< (\overline{x}_1-\overline{x}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}.$$

درجة الثقة تكون مضبوطة عندما تختار العينات من مجتمعات طبيعية. للمجتمعات الغير طبيعية يمكن الحصول على فترات ثقة تقريبية والتي تكون جيدة جدا عندما n_1 , n_2 تزيد عن 30. إذا كانت σ_1^2 , σ_2^2 مجهولتين والعينات المختارة كبيرة بدرجة كافية ، فإنه يمكن استبدال σ_1^2 , σ_2^2 بـ s_1^2 , s_2^2 على التوالي بدون التأثير على فترة الثقة .

مثال (٨-٤) أعطى اختبار في مادة الإحصاء إلى 75 طالبة و 50طالباً. فإذا كان متوسط النقاط من عينة الطالبات $\bar{x}_1 = 80$ بانحراف معياري $s_1=7$. وكان متوسط النقـــاط لعينـــة الطلبـــة $\bar{x}_1 = 70$ بانحراف معياري $s_2=6$. أوجد 95% فترة ثقة لـــ $\mu_1 - \mu_2$.

الحل . التقدير بنقطة لـــ $\mu_1 - \mu_2$ هو $\bar{x}_1 - \bar{x}_2 = 80-70=10$. وحيث أن كلا من n_1 , n_2 كبيرة فإنه يمكن استخدام $s_1=7$ بدلا من σ_1 و $s_2=6$ بدلا من σ_2. باستخدام $\alpha = 0.05$ فإن $z_{0.025} = 1.96$ وذلك من جدول التوزيع الطبيعي في ملحق (٣). وبالتعويض في 95% فترة ثقة لـــ $\mu_1 - \mu_2$ التالية :

$$(\bar{x}_1 - \bar{x}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2$$

$$< (\bar{x}_1 - \bar{x}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

نحصل على 95% فترة ثقة على الشكل :

$$10 - 1.96\sqrt{\frac{7^2}{75} + \frac{6^2}{50}} < \mu_1 - \mu_2 < 10 - 1.96\sqrt{\frac{7^2}{75} + \frac{6^2}{50}}$$

أو

$$7.703 < \mu_1 - \mu_2 < 12.297.$$

تستخدم الطريقة السابقة لتقدير الفرق بين متوسطين إذا كان σ_1^2 , σ_2^2 معلومتان أو يمكن تقديرهما من عينات كبيرة. إذا كانت أحجام العينات صغيرة ، لابد من استخدام توزيع t للحصول على فترات ثقة والتي تكون صحيحة عندما تكون المجتمعات تقريباً تتبع التوزيع الطبيعي.

بفرض أن $\sigma_1^2 = \sigma_2^2 = \sigma^2$ فإنه يمكن استخدام s_p^2 كتقدير للتباين العام σ^2. وباستخدام نظرية (٧-١٠) ، يتضح من شكل (٨-٤) أن :

$$P(-t_{\frac{\alpha}{2}} < T < t_{\frac{\alpha}{2}}) = 1 - \alpha.$$

حيث :

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\left(\frac{1}{n_1}\right) + \left(\frac{1}{n_2}\right)}}$$

و $t_{\frac{\alpha}{2}}$ هي قيمة t ، بدرجات حرية $\nu = n_1+n_2-2$ ، والتي المساحة على يمينها تساوي $\frac{\alpha}{2}$.

باستبدال T بقيمتها يمكن كتابة :

$$P\left[(\overline{X}_1-\overline{X}_2)-t_{\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < \mu_1-\mu_2 < (\overline{X}_1-\overline{X}_2)+t_{\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]=1-\alpha.$$

لأي عينتين عشوائيتين مستقلتين من الحجم n_1 , n_2 يتم اختيارهما من مجتمعين طبيعيين فإن الفرق بين متوسطي العينتين ، $\overline{x}_1-\overline{x}_2$ ، والتباين العام للعينة s_p^2 يتم حسابهما واستخدامهما في إيجاد $(1-\alpha)100\%$ فترة ثقة لـ $\mu_1-\mu_2$ على الشكل :

$$(\overline{x}_1-\overline{x}_2)-t_{\frac{\alpha}{2}}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < \mu_1-\mu_2 < (\overline{x}_1-\overline{x}_2)+t_{\frac{\alpha}{2}}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}.$$

مثال (٨-٥) اختيرت مجموعتان من الأرانب ، الأولى من 13 أرنباً وأعطيت الغذاء A والثانية من 15 أرنباً وأعطيت الغذاء B وكانت الزيادة في الوزن بعد فترة معينة هي :

A: 35, 30, 30, 23, 21, 12, 24, 23, 33, 27, 29, 25, 21.
B: 20, 17, 34, 31, 29, 39, 30, 46, 7, 21, 33, 43, 21, 34, 20.

أوجد 95% فترة ثقة للفرق بين متوسطي المجتمعين ، وذلك تحت فرض أن المجتمعين تقريباً يتبعان التوزيع الطبيعي حيث $\sigma_1^2 = \sigma_2^2$.

الحل .

$$n_1=13, \quad \overline{x}_1=25.62, \quad s_1=6.05,$$
$$n_2=15, \quad \overline{x}_2=28.33, \quad s_2=10.58,$$

لإيجاد فترة ثقة لـ $\mu_1-\mu_2$ سوف نستخدم التقدير بنقطة $\overline{x}_1-\overline{x}_2=25.62-28.33=-2.71$

. التباين العام s_p^2 هو :

$$s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

$$= \frac{(12)(6.05)^2 + (14)(10.58)^2}{13 + 15 - 2}$$
$$= 77.1669.$$

بأخذ الجذر التربيعي للتباين العام فإن s_p=8.784. باستخدام $\alpha = .025$ فـــإن $t_{.025}$=2.056 تستخرج من جدول توزيع t في ملحق (٤) عند درجات حريـــة $\nu = 13 + 15 - 2 = 26$.

بالتعويض في الصيغة التالية :

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2$$
$$< (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

فإن 95% فترة ثقة لـــ $\mu_1 - \mu_2$ هي :

$$-2.71 - (2.056)(8.784)\sqrt{\frac{1}{13} + \frac{1}{15}} < \mu_1 - \mu_2 <$$
$$-2.71 + (2.056)(8.784)\sqrt{\frac{1}{13} + \frac{1}{15}}.$$

والتي يمكن اختزالها الى :

$$-9.553 < \mu_1 - \mu_2 < 4.133.$$

تفترض الطريقة السابقة للحصول على فترات ثقة لـــ $\mu_1 - \mu_2$ أن المجتمعين طبيعين وأن $\sigma_1^2 = \sigma_2^2$. أيضا يمكن الحصول على نتائج جيدة إذا كانت $\sigma_1^2 \neq \sigma_2^2$ وذلك تحـــت شرط أن المجتمعين طبيعين و $n_1 = n_2$.

الآن وعند الرغبة في إيجاد $(1-\alpha)100\%$ فترة ثقة لـــ $\mu_1 - \mu_2$ في حالة العينـــات الصغيرة عندما تكون $\sigma_1^2 \neq \sigma_2^2$ وعند صعوبة الحصول على عينات ذات أحجام متساوية. الإحصاء الأكثر إستخداماً في هذه الحالة هو :

$$T' = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{S_1^2}{n_1}\right) + \left(\frac{S_2^2}{n_1}\right)}}$$

والذي تقريباً يتبع توزيع t بدرجات حرية v ، حيث :

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left[\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_1}\right)^2}{n_2 - 1}\right]}$$

وبما أن v نادراً ما تكون عدد صحيح ، فإننا نقربها الى أقرب رقم صحيح. وعلى ذلك يمكن كتابة :

$$P(-t_{\frac{\alpha}{2}} < T' < t_{\frac{\alpha}{2}}) \approx 1 - \alpha.$$

حيث $t_{\frac{\alpha}{2}}$ هي قيمة لتوزيع t ، بدرجات حرية v ، والتي المساحة على يمينها تساوى $\frac{\alpha}{2}$.

بالتعويض عن T' بقيمتها في المتباينة وإتباع الخطوات نفسها المتبعة في الحالات السابقة يمكن الحصول على $(1-\alpha)100\%$ فترة ثقة لـ $\mu_1 - \mu_2$ كالتالي :

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2$$

$$< (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} .$$

مثال (٨ –٦) خلال 20 سنة ماضية كان متوسط سقوط المطر في المنطقة A في قطر ما خلال شهر يناير 1.8 بوصة بانحراف معياري 0.4 بوصة. بينما كان متوسط سقوط المطر في المنطقة B من نفس القطر خلال 15 سنة ماضية 1.03 بوصة بانحراف 0.25 بوصة . أوجد 95% فترة ثقة لـ $\mu_1 - \mu_2$ وذلك تحت فرض أن المفردات مأخوذة من مجتمعين طبيعيين حيث $\sigma_1^2 \neq \sigma_2^2$.

الحل . المنطقة : A فإن $n_1 = 20, \quad \bar{x}_1 = 1.8, \quad s_1 = 0.4,$

وللمنطقة B فإن $n_2 = 15, \quad \bar{x}_2 = 1.03, \quad s_2 = 0.25,$

وعلى ذلك 95% فترة ثقة لـ $\mu_1 - \mu_2$ حيث $\sigma_1^2 \neq \sigma_2^2$ و $n_1 \neq n_2$ يمكـن الحصول عليها بالاعتماد على توزيع t بدرجة حرية تحسب كالآتي :

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left[\frac{(\frac{s_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{s_2^2}{n_1})^2}{n_2 - 1}\right]}$$

$$= \frac{\left(\frac{.4^2}{20} + \frac{.25^2}{15}\right)^2}{\left[\frac{\left(\frac{.4^2}{20}\right)^2}{19}\right] + \left[\frac{\left(\frac{.25^2}{15}\right)^2}{14}\right]} = 32.11 \approx 32.$$

وعلى ذلك $\bar{x}_1 - \bar{x}_2 = 1.8 - 1.03 = 0.77$ وتحت فرض أن $\alpha = 0.05$ ومن الجدول في ملحق (٤) فإن $t_{0.025} = 2.042$ بدرجات حرية $\nu = 32$. وبالتعويض في الصيغة التالية :

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \; < \; \mu_1 - \mu_2$$

$$< \; (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \; .$$

يمكن إيجاد 95% فترة ثقة (تقريبية) كالتالي :

$$0.77 - 2.042\sqrt{\frac{.4^2}{20} + \frac{.25^2}{15}} < \mu_1 - \mu_2 <$$

$$0.77 + 2.042\sqrt{\frac{.4^2}{20} + \frac{.25^2}{15}}.$$

والتي تختزل إلى :

$$\mathbf{0.545 < \mu_1 - \mu_2 < 0.995} \, .$$

وأخيراً سوف نناقش في هذا البند طريقة تقدير الفرق بين متوسطين عندما تكون العينتين غير مستقلين. فعلى سبيل المثال عندما تأخذ عينة واحدة ونحصل على قراءات لمفرداقا ثم نضع هذه العينة تحت مؤثر ونعود ونأخذ قراءات أخري لها ، وبمقارنة مجموعتي القراءتين لنفس المفردات يمكننا استنتاج تأثير هذا العامل أو المؤثر. لنفرض مثلا أننا نريد معرفة تأثير دواء على قراءات ضغط الدم المرتفع وأخذنا لذلك عينة من 10 شخصاً وقرأنا ضغط الدم لكل منهما ثم أعطينا كل شخص دواء له تأثير على ضغط الدم المرتفع وأعدنا أخذ القراءات مرة أخرى. في هذه الحالة نقول أننا أمام عينتين غير مستقلين أو عينتين مزدوجتين **paired samples**. أزواج المشاهدات سوف تكون $(x_1 - y_1),(x_2 - y_2),\ldots,(x_n - y_n)$ الفروق لازواج المشاهدات سوف تكون $d_1 = (x_1 - y_1), d_2 = (x_2 - y_2),\ldots, d_n = (x_n - y_n)$. هذه الفروق تمثل قيم للمتغير العشوائي D. التقدير بنقطة لمتوسط المجتمع $\mu_D = \mu_1 - \mu_2$ يعطى من $\bar{d}$ والذي يساوى متوسط الفروق في العينة. وبما أن $\bar{d}$ تمثل قيمة للإحصاء $\bar{D}$ كما أن التباين للفروق هو s_d^2 حيث :

$$s_d^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} d_i^2 - \frac{(\sum_{i=1}^{n} d_i)^2}{n}\right]$$

يمثل قيمة للمتغير العشوائي S_d^2 فإن $(1-\alpha)100\%$ فترة ثقة للمعلمة μ_D يمكن الحصول عليها باستخدام نظرية (٧-١١) والتي تسمح بكتابة :

$$P(-t_{\frac{\alpha}{2}} < T < t_{\frac{\alpha}{2}}) = 1-\alpha.$$

حيث :

$$T = \frac{\bar{D} - \mu_D}{\frac{S_d}{\sqrt{n}}}$$

و $t_{\frac{\alpha}{2}}$ هي قيمة لتوزيع t ، بدرجات حرية $v = n-1$ والتي تكون المساحة على يمينها تساوى $\frac{\alpha}{2}$. الآن سوف نتبع نفس الخطوات المتبعة في الحالات السابقة وذلك للحصول على $(1-\alpha)100\%$ فترة ثقة كالآتي :

$$\bar{d}-t_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}}<\mu_D<\bar{d}+t_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}}.$$

حيث $\bar{d}$ و s_d هما المتوسط والانحراف المعياري للفروق لعدد n من ازواج المشاهدات و $t_{\frac{\alpha}{2}}$

هي قيمة لتوزيع t بدرجات حرية $v = n-1$ والتي تكون المساحة على يمينها تساوى $\frac{\alpha}{2}$.

مثال (٨-٧) أخذت عينة عشوائية من 10 تلاميذ من إحـدى المدارس ودونـت أوزانهـم ثم أعطي كل منهم كوباً من اللبن صباحاً وآخر ظهراً وذلك لمدة ثلاثة شهور متتاليـة . ثم دونـت أوزانهم فكانت النتائج كالآتي :

الوزن قبل تعاطي اللبن	129	124	126	139	133	136	139	135	137	140
الوزن بعد تعاطي اللبن	130	126	129	140	136	134	141	140	138	141

المطلوب إيجاد 99% فترة ثقة للفرق الحقيقي $\mu_D=\mu_1-\mu_2$.

الحل . التقدير بنقطة لـ μ_D هو $\bar{d}=-1.7$. التباين s_d^2 لفروق العينة هو :

$$s_d^2=\frac{1}{n-1}\left[\Sigma d_i^2-\frac{(\Sigma d_i)^2}{n}\right]$$

$$=\frac{1}{9}\left[59-\frac{(-17)^2}{10}\right]=3.344.$$

وبأخذ الجذر التربيعي للمقدار s_d^2 فإن $s_d=1.829$. باستخدام $\alpha=0.01$ فإن $t_{0.005}=3.25$ والمستخرجة من جدول توزيع t في ملحق (٤) عند درجات حرية $v=n-1=9$. وبالتعويض في الصيغة التالية :

$$\bar{d}-t_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}}<\mu_D<\bar{d}+t_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}}.$$

نحصل على 99% فترة ثقة كالتالي :

$$-1.7-(3.25)\frac{(1.829)}{\sqrt{10}}<\mu_D<-1.7+(3.25)\frac{(1.829)}{\sqrt{10}}$$

والتي تختزل الى :

$$-3.58<\mu_D<0.18.$$

(٨-٤) فترة ثقة للنسبة Confidence Interval for Proportion

أوضحنا في الفصل السابع أن الإحصاء $\overline{X}$ ليس دائما المطلوب في مجـــال الإحصـــاء. ففي بعض الأحيان يكون الاهتمام بمعرفة نسبة وجود صفة معينة في مجتمع ما مثل نسبة المصـــابين بتسوس الأسنان أو نسبة النباتات المصابة وهكذا. التقدير بنقطة الأكثر كفاءة لنسبة صفة ما p هو الإحصاء $\hat{P} = \frac{X}{n}$. وعلى ذلك ، فإن نسبة العينة $\hat{p} = \frac{x}{n}$ سوف تستخدم كتقدير بنقطـــة للمعلمة p. من الفصل السابق ومن نظرية (٧-٧) نعلم أن $\hat{p}$ تقريباً تتبـــع توزيعـــاً طبيعيـــاً متوسطة $\mu_{\hat{p}} = p$ وتباينه $\sigma^2_{\hat{p}} = \frac{pq}{n}$. وعلى ذلك فإن :

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1-\alpha \ .$$

حيث :

$$Z = \frac{\hat{P}-p}{\sqrt{\frac{pq}{n}}}$$

و $z_{\frac{\alpha}{2}}$ هي قيمة للمتغير العشوائي الطبيعي القياسي Z والتي المساحة على يمينها تساوى $\frac{\alpha}{2}$.

بالتعويض عن Z بقيمتها فإنه يمكن كتابة :

$$P(-z_{\frac{\alpha}{2}} < \frac{\hat{P}-p}{\sqrt{\frac{pq}{n}}} < z_{\frac{\alpha}{2}}) = 1-\alpha \ .$$

وبضرب كل حد في المتباينة في $\sqrt{\frac{pq}{n}}$ وطرح $\hat{p}$ والضرب في 1- ، نحصل على :

$$P(\hat{P} - z_{\frac{\alpha}{2}}\sqrt{\frac{pq}{n}} < p < \hat{P} + z_{\frac{\alpha}{2}}\sqrt{\frac{pq}{n}}) = 1-\alpha \ .$$

وحيث أن p عادة ما تكون مجهولة لذلك تستخدم $\hat{p}$ بدلا من p وعلى ذلك فإن :

$$P(\hat{P} - z_{\frac{\alpha}{2}} < \sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{P} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}) \simeq 1-\alpha \ .$$

لعينة عشوائية خاصة من الحجم n (تبعا لقواعد Cochran) كما ذكرنا في الفصـــل الســـابع نحسب نسبة العينة $\hat{p} = \frac{x}{n}$ ويتم الحصول على $(1-\alpha)100\%$ فترة ثقة للمعلمة p كمايلي :

$$\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} .$$

مثال (٨-٨) يقوم مصنع بإنتاج منتج على درجة عالية من الجودة ويرغب المسئول في المصنـــع في تقدير نسبة الوحدات المنتجة التالفة. فإذا اختيرت عينة عشوائية من 200 وحدة ووجـــد أن بينهم 40 وحدة تالفة ، أوجد 95% فترة ثقة للمعلمة p.

الحل . التقدير بنقطة للمعلمة p هو $\hat{p} = \frac{40}{200} = 0.2$. باستخدام جدول التوزيع الطبيعـــي في ملحق (٣) فإن $z_{0.025} = 1.96$. بالتعويض في الصيغة التالية :

$$\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

يمكن الحصول على 95% فترة ثقة كالتالي :

$$0.2 - 1.96\sqrt{\frac{(0.2)(0.8)}{200}} < p < 0.2 + 1.96\sqrt{\frac{(0.2)(0.8)}{200}}$$

والتي يمكن اختزالها الى :

$$0.145 < p < 0.255 .$$

إذا وقعت p عند مركز $(1-\alpha)100\%$ فترة الثقة ، فإن $\hat{p}$ سوف تقـــدر p بـــدون أخطاء. في معظم الأحوال ، $\hat{p}$ لا تساوي p وعلى ذلك يكون هناك فرق بين $\hat{p}$ و p والـــذي يمثل الخطأ. هذا الخطأ يصل الى أقصاه عندما تكون p قريبة من إحدى حديّ الثقة. وعلى ذلك $\hat{p}$ سوف تختلف عن p بقيمة أقل من $z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$ كما يتضح من شكل (٨-٥).

Error

$\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$ $\hat{p}$ p $\hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$

شكل (٨-٥)

نظرية (٨-٣) إذا استخدمت $\hat{p}$ كتقدير للمعلمة p فإنه يكون لدينا $(1-\alpha)100\%$ ثقة أن الخطأ سوف يكون أقل من $z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$.

عند الرغبة في تقدير حجم العينة اللازم للتأكد من أن الخطأ في تقدير p اقل من مقدار معين e وتبعا لنظرية (٨-٣) فلا بد من اختيار n بحيث تكون $e = z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$.

نظرية (٨-٤) إذا استخدمت $\hat{p}$ كتقدير بنقطة للمعلمة p فإنه يكون لدينا $(1-\alpha)100\%$ ثقة أن الخطأ سوف يكون أقل من قيمة معينة e عندما يحسب حجم العينة من الصيغة التالية :

$$n = \frac{z^2_{\frac{\alpha}{2}}\ \hat{p}\hat{q}}{e^2}$$

وحيث أن $\hat{p}$ في الصيغة السابقة لا بد أن تقدر من عينة ، لذلك لابد من اختيار عينة مبدئية كبيرة وحساب نسبة العينة $\hat{p}$ منها.

مثال (٨-٩) أحسب القيمة العظمى للخطأ في التقدير e عند 95% للمثال (٨-٨).

الحل . حيث e تعطي بالعلاقة :

$$\mathbf{e = z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}}$$

وحيث أن $z_{\frac{\alpha}{2}} = 1.96$, $\hat{p} = 0.2$, $\hat{q} = 0.8$ فإن:

$$e = 1.96\sqrt{\frac{(0.2)(0.8)}{200}} = 0.0554.$$

مثال (٨-١٠) في عينة عشوائية من 500 مواطن في مجتمع سكاني ما ، وجد منهم 270 مواطناً يحبون أن يضاف الى مياههم قليل من الفلور. المطلوب :

(أ) إيجاد 95% فترة ثقة لنسبة المجتمع الذين يحبذون إضافة الفلور.

(ب) تقدير حجم العينة التي يمكننا التأكد منها باحتمال 95% من أن الخطأ لا يتجاوز 0.05.

الحل . (أ) التقدير بنقطة للمعلمة p هو $\hat{p} = \frac{x}{n} = \frac{270}{500} = 0.54$. باستخدام جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن $z_{0.025} = 1.96$ وبالتعويض في الصيغة التالية :

$$\hat{p}-z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}<p<\hat{p}+z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

يمكن الحصول على 95% فترة ثقة كالتالي :

$$0.54-1.96\sqrt{\frac{(0.54)(0.46)}{500}}<p<0.54+1.96\sqrt{\frac{(0.54)(0.46)}{500}}$$

والتي تختزل إلى :

$$\mathbf{0.496 < p < 0.584.}$$

(ب) باعتبار الأشخاص الذين عددهم 500 يمثلون عينة عشوائية مبدئية حيث أن $\hat{p}=0.54$ وباستخدام نظرية (٨-٤) فإن :

$$n=\frac{(1.96)^2(0.54)(0.46)}{(0.05)^2}=381.70$$

$$\approx 382.$$

(٨-٥) فترة ثقة للفرق بين نسبتين

Confidence Interval for the Difference Between Two Proportions

المقدر بنقطة الأكثر كفاءة للفرق بين نسبتين ، p_1-p_2 ، هو الإحصاء $\hat{P}_1-\hat{P}_2$. وعلى ذلك ، للحصول على تقدير بنقطة لـ p_1-p_2 سوف تختار عينتين عشوائيتين مستقلتين من الحجم n_2, n_1 وحساب النسبة للصفة موضع الدراسة في كل عينة ، $\hat{p}_1=\frac{x_1}{n_1}, \hat{p}_2=\frac{x_2}{n_2}$، حيث x_2 , x_1 يمثلان عدد المفردات الذين يملكون الصفة موضع الاهتمام في العينتين على التوالي. يتم حساب الفرق $\hat{p}_1-\hat{p}_2$. فترة ثقة لـ $\hat{P}_1-\hat{P}_2$ يمكن الحصول عليها بالاعتماد على التوزيع العيني للإحصاء $\hat{P}_1-\hat{P}_2$ والذي تقريباً ، تبعا لنظرية (٧-٨) ، يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{p}_1-\hat{p}_2}=\mu_1-\mu_2$$

وانحراف معياري :

$$\sigma_{\hat{p}_1-\hat{p}_2}=\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}.$$

وعلى ذلك يمكن كتابة :

$$P(-z_{\frac{\alpha}{2}}<Z<z_{\frac{\alpha}{2}})=1-\alpha,$$

حيث أن :

$$Z=\frac{(\hat{P}_1-\hat{P}_2)-(p_1-p_2)}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}}$$

وحيث $z_{\frac{\alpha}{2}}$ هي قيمة لمنحنى التوزيع الطبيعي القياسي والتي تكون المساحة على يمينها تساوى $\frac{\alpha}{2}$. بالتعويض عن Z بقيمتها يمكن كتابة :

$$P\left[-z_{\frac{\alpha}{2}}<\frac{(\hat{P}_1-\hat{P}_2)-(p_1-p_2)}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}}<z_{\frac{\alpha}{2}}\right]=1-\alpha .$$

وبضرب كل حد في المتباينة في $\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}$ وطرح $\hat{P}_1-\hat{P}_2$ والضرب في 1- نحصل على :

$$P\left[(\hat{P}_1-\hat{P}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}<p_1-p_2 <(\hat{P}_1-\hat{P}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}\right]=1-\alpha.$$

إذا كانت كلا من n_2, n_1 كبيرة فإنه يمكن استبدال $\hat{p}_1-\hat{p}_2$ بتقديراتهما $\hat{p}_1=\frac{x_1}{n_1}, \hat{p}_2=\frac{x_2}{n_2}$ وعلى ذلك فإن :

$$P\left[(\hat{P}_1-\hat{P}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}<(p_1-p_2) <(\hat{P}_1-\hat{P}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}\right]\approx 1-\alpha .$$

وبالتالي يمكن الحصول على $(1-\alpha)100\%$ فترة ثقة للفرق بين نسبتين كالتالي :

$$(\hat{p}_1-\hat{p}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}<(p_1-p_2)$$

$$<(\hat{p}_1-\hat{p}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}.$$

مثال (٨ – ١١) في عينة من 2000 من الرجال و 5000 من النساء الذين يشاهدن برنامجا تليفزيونيا يوميا وجد أن 1100 من الرجال و 2300 من النساء يفضلون هذا البرنامج. أوجد 95% فترة ثقة للفرق بين نسبة كل من الرجال ونسبة كل من النساء الذين يشاهدون هذا البرنامج ويفضلونه.

الحل . بفرض أن p_1-p_2 النسبتين الحقيقتين وعلى ذلك

$$\hat{p}_2=\frac{2300}{5000}=0.46 \quad , \quad \hat{p}_1=\frac{1100}{2000}=0.55$$

وعلى ذلك التقدير بنقطة لـ p_1-p_2 هو $\hat{p}_1-\hat{p}_2=0.55-0.46=.09$. باستخدام جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن $z_{0.025}=1.96$ وبالتعويض في الصيغة التالية :

$$(\hat{p}_1-\hat{p}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}<p_1-p_2$$

$$<(\hat{p}_1-\hat{p}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}.$$

فإن :

$$0.09-1.96\sqrt{\frac{(0.55)(0.45)}{2000}+\frac{(0.46)(0.54)}{5000}}<p_1-p_2$$

$$<0.09+1.96\sqrt{\frac{(0.55)(0.45)}{2000}+\frac{(0.46)(0.54)}{5000}},$$

والتي تختزل الى :

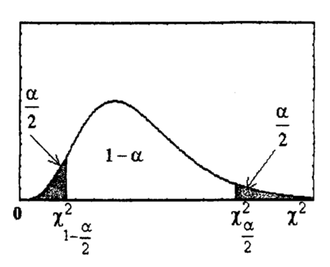
$$\mathbf{0.0642<p_1-p_2<0.1158.}$$

(٨–٦) فترة ثقة للتباين Confidence Interval for the Variance

التقدير بنقطة لتباين المجتمع σ^2 نحصل عليه من تباين العينة. وعلى ذلك يعتبر S^2 مقدر للمعلمة σ^2. التقدير بفترة للمعلمة σ^2 يمكن الحصول عليه باستخدام الإحصاء :

$$X^2 = \frac{(n-1)S^2}{\sigma^2}$$

وبالرجوع إلى نظرية (٧-١٢) فإن الإحصاء X^2 له توزيع χ^2 بدرجات حرية n-1 وذلك عندما تختار العينة من توزيع طبيعي. وبالرجوع الى شكل (٨-٦).



شكل (٨-٦)

يمكن كتابة :

$$P\left[\chi^2_{1-\frac{\alpha}{2}} < X^2 < \chi^2_{\frac{\alpha}{2}}\right] = 1-\alpha.$$

حيث $\chi^2_{\frac{\alpha}{2}}$ و $\chi^2_{1-\frac{\alpha}{2}}$ هما قيمتان لتوزيع χ^2 بدرجات حرية n-1 والتي المساحة على يمين $\chi^2_{\frac{\alpha}{2}}$ تساوى $\frac{\alpha}{2}$، والمساحة على يمين $\chi^2_{1-\frac{\alpha}{2}}$ تساوي $1-\frac{\alpha}{2}$. بالتعويض عن X^2 بقيمتها يمكن كتابة :

$$P\left[\chi^2_{1-\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}\right] = 1-\alpha.$$

بقسمة كل حد في المتباينة على S^2 (n-1) ، وعكس كل حد نحصل على :

$$P\left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}}\right] = 1-\alpha.$$

لعينة عشوائية خاصة من الحجم n ، فإن تباين العينة s^2 يحسب ويمكن الحصول على $(1-\alpha)100\%$ فترة ثقة للمعلمة σ^2 كما يلي :

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}.$$

مثال (٨-١٢) تسلم أحد التجار كمية كبيرة من بطاريات السيارات المنتجة بواسطة مصنع جديد وتم اختيار عينة عشوائية من البطاريات التي تسلمها التاجر وتمت تجربتها فكانت أعمارها بالشهر هي :

26.9	28.5	33.6	28.0	23.9
28.7	29.3	29.1	35.9	35.2

أوجد 99% فترة ثقة للمعلمة σ^2 .

الحل . أولا نحصل على تباين العينة s^2 وهو :

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}\right]$$

$$= \frac{1}{9}\left[9076.87 - \frac{(299.1)^2}{10}\right] = 14.53.$$

باستخدام جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $v = n-1 = 9$ فإن :

$$\chi^2_{0.995} = 1.735 \quad , \quad \chi^2_{0.005} = 23.587$$

بالتعويض في الصيغة التالية :

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}$$

يمكن الحصول على فترة ثقة كالتالي :

$$\frac{(9)(14.53)}{23.587} < \sigma^2 < \frac{(9)(14.53)}{1.735}$$

والتي يمكن اختزالها الى

$$5.544 < \sigma^2 < 75.372.$$

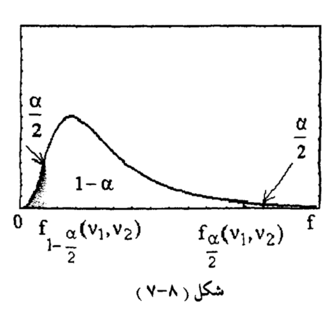
(٨-٧) فترة ثقة لنسبة تباينين

Confidence Interval for the Ratio of two variances

التقدير بنقطة لنسبة تبايني مجتمعين ، σ_1^2 / σ_2^2 ، يمكن الحصول عليه من النسبة ، s_1^2 / s_2^2 ، لتبايني عينتين. وعلى ذلك يعتبر S_1^2 / S_2^2 مقدر للنسبة σ_1^2 / σ_2^2. فإذا كان لدينا مجتمعان الأول يتبع توزيعاً طبيعياً متوسطة μ_1 وتباينه σ_1^2 والثاني يتبع توزيعاً طبيعياً متوسطة μ_2 وتباينه σ_2^2، فإنه يمكن الحصول على $(1-\alpha)100\%$ فترة ثقة لـ σ_1^2 / σ_2^2 باستخدام الإحصاء :

$$F = \frac{\sigma_2^2 \; S_1^2}{\sigma_1^2 \; S_2^2}.$$

تبعا لنظرية (٧-١٣) فإن المتغير العشوائي F له توزيع F بدرجات حرية $v_1 = n_1 - 1$, $v_2 = n_2 - 1$. وعلى ذلك يمكن كتابة (كما يتضح من شكل (٨-٧).



شكل (٨-٧)

$$P\left[f_{1-\frac{\alpha}{2}}(\nu_1,\nu_2) < F < f_{\frac{\alpha}{2}}(\nu_1,\nu_2)\right] = 1-\alpha.$$

حيث $f_{1-\frac{\alpha}{2}}(\nu_1,\nu_2)$, $f_{\frac{\alpha}{2}}(\nu_1,\nu_2)$ هما قيمتان لتوزيع F بدرجات حرية ν_2 , ν_1 على التوالي (كما يتضح من شكل (٨-٧)).وبالتعويض عن F بقيمتها فإن :

$$P\left[f_{1-\frac{\alpha}{2}}(\nu_1,\nu_2) < \frac{\sigma_2^2\ S_1^2}{\sigma_1^2\ S_2^2} < f_{\frac{\alpha}{2}}(\nu_1,\nu_2)\right] = 1-\alpha.$$

بضرب كل حد في المتباينة في S_2^2/S_1^2 وعكس كل حد نحصل على :

$$P\left[\frac{S_1^2}{S_2^2}\frac{1}{f_{\frac{\alpha}{2}(\nu_1,\nu_2)}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2}\frac{1}{f_{1-\frac{\alpha}{2}}(\nu_1,\nu_2)}\right] = 1-\alpha.$$

نتائج نظرية (٧-١٤) سوف تساعدنا في استبدال $f_{1-\frac{\alpha}{2}}(\nu_1,\nu_2)$ بـــ $f_{\frac{\alpha}{2}}(\nu_2,\nu_1)$

وعلى ذلك :

$$P\left[\frac{S_1^2}{S_2^2}\frac{1}{f_{\frac{\alpha}{2}(\nu_1,\nu_2)}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} f_{\frac{\alpha}{2}(\nu_2,\nu_1)}\right] = 1-\alpha.$$

لأي عينتين عشوائيتين مستقلتين من الحجم n_2, n_1 مأخوذتين من مجتمعين طبيعيـــين ، فإن النسبة s_1^2/s_2^2 تحسب ويتم الحصول على $(1-\alpha)100\%$ فترة ثقة كما يلي :

$$\frac{s_1^2}{s_2^2}\frac{1}{f_{\frac{\alpha}{2}(\nu_1,\nu_2)}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\frac{\alpha}{2}(\nu_2,\nu_1)}.$$

مثال (٨-١٣) إذا كانت درجات كل من الطلاب والطالبات بـــإحدى الجامعـــات في مـــادة الإحصاء يتبع توزيعاً طبيعياً. اختير عينة عشوائية من بين الطلاب وأخرى مـــن بـــين الطالبـــات فكانت درجاتهم كما يلي :

: الطلاب	59	79	73	49	88	83	69	44	81
: الطالبات		74	79	49	59	82	69	79	89

أوجد 90% فترة ثقة للنسبة σ_1^2 / σ_2^2 .

الحل .

$$n_1 = 9 \ , \ s_1 = 15.57 \ , \ n_2 = 8 \ , \ s_2 = 13.07$$

$f_{.05}(8,7) = 3.73$, $f_{.05}(7,8)=3.5$ المستخرجتان من جدول توزيع F في ملحق (٦) بدرجات حرية $v_1 = 8, \ v_2 = 7$ للعينة الأولي ودرجات حرية $v_1 = 7, \ v_2 = 8$ للعينة الثانية. يمكن الحصول على 90% فترة ثقة للنسبة σ_1^2 / σ_2^2 وذلك بالتعويض في الصيغة التالية

$$\frac{s_1^2}{s_2^2} \frac{1}{f_{\frac{\alpha}{2}}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\frac{\alpha}{2}}(v_2, v_1) \ .$$

أي أن :

$$\frac{(15.57)^2}{(13.07)^2 (3.73)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{(15.57^2)(3.5)}{(13.07)^2}$$

والتي تختزل إلى :

$$0.3805 < \frac{\sigma_1^2}{\sigma_2^2} < 4.967.$$

تمـــارين :

-١- إذا كان عمر المصابيح الكهربائية المنتجة في مصنع ما ، تقريباً تتبع التوزيع الطبيعي بانحراف معياري 40 ساعة . اختيرت عينة عشوائية من 25 مصباحاً ووجد أن متوسط العمر 780 ساعة. أوجد 99% فترة ثقة لمتوسط المجتمع لكل المصابيح المنتجة من هذا المصنع .

-٢- في مركز تجاري كبير ، صممت ماكينة للعصير بحيث أن كمية العصير المستخرجة منـــــها لكل كوب يتبع تقريباً توزيعاً طبيعياً بانحراف معياري 0.5 أوقية لكل كوب . أوجد 95% فـترة ثقة للمتوسط μ إذا اختيرت عينة عشوائية من 26 كوب ووجد أن $\bar{x} = 7.4$ أوقية.

-٣- إذا كانت الأطوال لعينة عشوائية من 50 طالباً لها متوسط $\bar{x} = 68.5$ بوصة والانحراف معياري $s = 2.7$ بوصة. المطلوب

(أ) إيجاد 99% فترة ثقة للمتوسط μ .

(ب) تقدير حجم العينة n حتى يمكننا التأكد باحتمال 0.99 من أن الخطأ لا يتجاوز 0.5 .

-٤- في دراسة عن تلوث الهواء بأكسيد الكبريت المنبعث من إحدى المصانع. اختيرت عينة عشوائية من قراءات 70 يومياً وحسب متوسط العينة فوجد أن يساوى 19.1 طناً بانحراف معياري قدره 5.22 أطنان . أوجد 95% فترة ثقة للمتوسط μ .

-٥- يرغب صاحب مصنع في تقدير حجم العينة n حتى يمكنه التأكد تأكداً معقولاً من أن تقديره وباحتمال 0.99 لن يكون مخطئاً بأكثر من 4 وحدات معينة إذا علم أن الانحراف المعياري يساوى 18 وحدة. أوجد حجم العينة التي تحقق الشروط التي وضعها صاحب المصنع.

-٦- مونت 35 سيارة بكمية قدرها 4.8 جالون من البنزين المضاف إليه مادة معينة وسارت حتى نفذ الوقود. وبعد تمام التجربة حسب متوسط الأميال التي قطعتها كل سيارة بكل جالون فكان $\bar{x} = 18.9$ لكل جالون بانحراف معياري $s = 2.4$. أوجد 95% فترة ثقة للمتوسط μ.

-٧- مجموعة قوامها 50 من حيوانات التجارب أطعمت نوعاً معيناً من المقننات الغذائية لمدة شهر فكان متوسط الزيادة في الوزن $\bar{x} = 30$ بانحراف معياري $s=5$ أوقية أوجد 99% فترة ثقة للمتوسط μ .

-٨- دلت الخبرة مع العمال المشتغلين في صناعة معينة أن الزمن الذي يحتاج إليه العامل لإكمال عمل يتبع توزيعاً طبيعياً بانحراف معياري 20 دقيقة فإذا اختيرت عينة عشوائية من 25 عاملاً ووجد أن $\bar{x} = 13$ دقائق أوجد 95% فترة ثقة للمتوسط μ .

-٩- إذا كان ضغط الدم لمجموعة من الأفراد يتبع التوزيع الطبيعي بانحراف معياري $\sigma = 2$ اختيرت عينة عشوائية حجمها 25 فردا من هذه المجموعة وكان متوسط ضغط الدم من العينة $\bar{x} = 91$ أوجد 99% فترة ثقة للمتوسط μ.

-١٠- اختيرت عينة عشوائية من 100 أسرة فكان متوسط دخل الأسرة في العينة $\bar{x} = 1000$ جنيها بانحراف معياري $s = 100$ جنيها. أوجد 95% فترة ثقة للمتوسط μ.

-١١- اختيرت عينة عشوائية من 100 مريض بالسكر وكان متوسط أعمارهم $\bar{x} = 55$ بانحراف معياري $s = 20$ أوجد 99% فترة ثقة للمتوسط μ.

-١٢- لدراسة النمو لنوع معين من الزهور اختيرت عينة عشوائية من 70 زهرة ووجد أن متوسط النمو خلال عام $\bar{x} = 40.9$ بانحراف معياري 5.1 . أوجد 95% فترة ثقة للمتوسط μ.

-١٣- في بحث ميداني ، اختيرت عينة عشوائية من 400 عائلة ووجد أن متوسط ما أنفق على الطعام خلال عام هو 1200 جنية بانحراف معياري قدرة 100 جنيها أوجد 99% فترة ثقة للمتوسط μ.

-١٤- بفرض أن أوزان الدببة في لعب الأطفال يتبع التوزيع الطبيعي بانحراف معياري 3.5 كيلو جرام ما هو حجم العينة اللازم بثقة 95% أن متوسط العينة لا يختلف عن متوسط المجتمع بأكثر من 0.5 كيلو جرام .

-١٥- إذا كان أطوال الطلبة المستجدين في كلية ما يتبع توزيعاً طبيعياً بانحراف معياري 2.5 بوصة. فإذا اختيرت عينة عشوائية من الحجم $n = 25$ من هؤلاء الطلبة ووجد أن $\bar{x} = 68.5$ بوصة أوجد 95% فترة ثقة للمتوسط μ.

-١٦- في دراسة عن حجم المبيعات اليومية في محل تجارى تم تسجيل حجم المبيعات خلال 60 يوماً ، خلال فترة قدرها سنة ، وتم حساب متوسط المبيعات اليومية فكانت $\bar{x} = 218$ جنيها بانحراف معياري $s = 18$ جنيها. المطلوب تقدير 99% فترة ثقة للمتوسط μ.

-١٧- إذا كانت أسعار إحدى السلع تتبع توزيعاً طبيعياً بانحراف معياري 4 جنيها. اختيرت عينة عشوائية من أسعار هذه السلعة حجمها $n = 15$ فكان وسطها الحسابي $\bar{x} = 40$ جنيها. أوجد 99% , 95% فترة ثقة للمتوسط μ.

-١٨- إذا كان معروفا أن الضغط الداخلي لكرات التنس المنتجة بواسطة أحد المصانع يتبع توزيعاً طبيعياً بانحراف معياري $\sigma = 0.25$ رطلاً لكل بوصة. اختيرت عينة عشوائية من الحجم $n = 28$ من الكرات المنتجة من هذا المصنع وكان $\bar{x} = 28$ رطلاً لكل بوصة. أوجد 99% فترة ثقة للمتوسط μ.

-١٩- اختيرت عينة عشوائية من الحجم $n=50$ من إنتاج آلة ما لتعبئة الأرز في أكياس فحصلنا على البيانات التالية $\bar{x} = 4.80$ كجم و $s = 0.6$ كجم. أوجد 95% فترة ثقة للمتوسط μ.

-٢٠- اختيرت عينة عشوائية من 100 فاتورة مباعة وذلك من مجتمع كبير جداً من الفواتير المباعة. فإذا كان متوسط العينة هو $\bar{x} = 18.5$ جنيها بانحراف معياري 6.0 جنيها. أوجد 99% فترة ثقة للمتوسط μ.

-٢١- اختيرت عينة عشوائية من الفواتير المباعة من الحجم $n = 100$ وتم وضعها في التوزيع التكراري التالي :

251-300	201-250	151-200	101-150	51-100	1-50	حدود الفئة

←

عـــدد الفواتير	10	19	15	26	13	17

استخدم المعلومات في التوزيع التكراري السابقة في إيجاد 95% و 99% فترة ثقة لمتوسط المجتمع.

-٢٢- إذا كانت أوزان 10 صناديق من القمح هي :

12, 10.2, 10.1, 9, 10.1, 10.2, 10.1, 9.8, 9.9, 8.9

(الأوزان مقاسه بالأوقية). أوجد 95% فترة ثقة للمتوسط μ وذلك تحت فرض أن وزن الصندوق من القمح يتبع توزيعاً طبيعياً .

-٢٣- لدراسة اختفاء العلامات الأرضية البيضاء وسط الطرق نتيجة المرور الكثيف. أخذت عينة من 8 مناطق مختلفة ولوحظ اختفاء العلامات بعد مرور السيارات مقربة لأقرب مائة سيارة كالتالي :

163800, 136400, 108200, 125400,
143700, 163000, 159400, 122600

أوجد 99% فترة ثقة كالمتوسط μ وذلك تحت فرض أن توزيع المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً .

-٢٤- أخذت عينة عشوائية من 20 آلة حاسبة من نوع ما ووجد أن متوسط أعمارهم $\bar{x} = 8$ سنوات بانحراف معياري s=2.6 سنة . أوجد 95% فترة ثقة للمتوسط μ تحت فرض أن توزيع المجتمع الذي اختيرت منه العينة طبيعياً .

-٢٥- إذا كانت كمية النيكوتين في 7 سجائر من نوع معين مقاس بالمليجرامات هو

21, 19, 23, 19, 23, 18, 17

أوجد 95% فترة ثقة للمتوسط μ تحت فرض أن كمية النيكوتين في هذا النوع من السجائر تتبع توزيعاً طبيعياً.

-٢٦- في عينة عشوائية من 10 مرضى في العناية المركزة في مستشفي ما وجد أن متوسط درجة حرارة الجسم 38.1 درجة بانحراف معياري 1.2 . أوجد 95% فترة ثقة لمتوسط المجتمع μ تحت فرض أن متوسط درجة حرارة الجسم للمرضى في العناية المركز يتبع توزيعاً طبيعياً.

-٢٧- في إحدى المراكز الصحية أجريت دراسة على عينة مبدئية ، على الكمية المتوقعة من الأكسجين (باللتر في الدقيقة) الذي يستهلكه الفرد الذي عمره 25-20 سنة ووجد أن الانحراف المعياري 0.5 لتر في الدقيقة. المطلوب تقدير حجم العينة اللازم لبحث قادم حتى يمكننا *التأكد باحتمال 0.99 من أن الخطأ لا يتجاوز 0.1 لتر في الدقيقة .*

-٢٨- إذا كانت أوزان طلاب إحدى المدارس تتبع توزيعاً طبيعياً. اختيرت عينة عشوائية حجمها 16 طالباً من هذه المدرسة فكانت أوزانهم بالكيلو جرام كما يلي :

42.1	44.2	49	44	45	44	51	52
45.2	46	51.3	44	47	49	41.2	47

أوجد 95% فترة ثقة للمتوسط μ.

-٢٩- في تجربة على 10 من رواد الفضاء في مجال يحاكي مجال انعدام الوزن وجد أن متوسط ضربات القلب لهم 26.22 دقة في الدقيقة بانحراف معياري 4.3 دقة في الدقيقة. أوجد 99% فترة ثقة للمتوسط μ تحت فرض أن ضربات القلب لرائد الفضاء تتبع توزيعاً طبيعياً .

-٣٠- في دراسة للسعرات الحرارية المنتجة من نوع معين من الفحم تم الحصول على البيانات التالية من عينة عشوائية حجمها $n=5$:

8000, 7820, 8200, 8470, 8123

أوجد 95% فترة ثقة للمتوسط μ تحت فرض أن السعرات تحت الدراسة تتبع توزيعاً طبيعياً .

-٣١- اختيرت عينة عشوائية من 10 كرات وتم قياس قطر كل كرة وحساب متوسط القطر فكان $\bar{x} = 4.9$ ملم بانحراف معياري 0.07 ملم. أوجد 95% فترة ثقة للمتوسط μ (تحت فرض أن العينة تم اختيارها من مجتمع طبيعي).

-٣٢- لمقارنة صنفين من القمح من حيث كمية المحصول أخذت 5 فدادين لكل صنف من القمح وزرع فيها الصنفين تحت نفس الظروف. أعطى الصنف A في المتوسط 78.3 وحدة ما لكل فدان بانحراف معياري 5.5 وحدة لكل فدان. بينما أعطى الصنف B في المتوسط 88 وحدة لكل فدان بانحراف معياري 6.2 أوجد 95% فترة ثقة لـ $\mu_1 - \mu_2$ تحت فرض أن العينتين تم اختيارهما من مجتمعين طبيعيين.

-٣٣- في دراسة لتقدير الفرق بين الأجور في كليتين A و B أ أخذت عينة عشوائية من $n_1=25$ أستاذ من الكلية A ووجد أن متوسط الأجر خلال 9 شهور هو 10000دولار بانحراف معياري 1100 دولار. كما أخذت عينة عشوائية أخرى من الحجم $n_2=20$ أستاذ من الكلية B ووجد أن متوسط الأجر خلال 9 شهور 13000 دولار بانحراف معياري 1120 دولار. أوجد 95% فترة ثقة لـ $\mu_1 - \mu_2$ تحت فرض أن $\sigma_1^2 = \sigma_2^2$ وأن العينتين تم اختيارهما من توزيعين طبيعيين.

-٣٤- تحاول شركة لتأجير السيارات اتخاذ قرار بشأن شراء إطارات من النوع A أو من النوع B. لتقدير الفرق بين النوعين أجرت تجربة حيث استخدمت 12 إطاراً من كل نوع وجربت الإطارات حتى انتهاء عمرها وكانت النتائج كما يلي (المسافة التي قطعتها السيارة بالأميال):

A النوع الأول : $\bar{x}_1 = 22500$, $s_1 = 310$

B النوع الثاني : $\bar{x}_2 = 23600$, $s_2 = 380$

بفرض أن العينتين تتم اختيارهما من توزيعين طبيعيين وأن $\sigma_1^2 \neq \sigma_2^2$ أوجد 95% فترة ثقة لـــ $\mu_1 - \mu_2$.

-٣٥- البيانات التالية تمثل الزمن اللازم لعرض فيلم منتج من قبل شركتين مختلفتين .

	الزمن (بالدقائق)						
الشركة الأولى	100	90	110	80	90		
الشركة الثانية	96	120	90	174	80	108	107

أحسب 90% فترة ثقة لـــ $\mu_1 - \mu_2$ بفرض أن أزمنة العرض تقريباً تتبع التوزيع الطبيعي.

-٣٦- أجريت دراسة على طريقة جديدة لتخفيض الوزن 10 أرطال في المتوسط على 7 سيدات تابعوا نفس الطريقة لتخفيض الوزن وسجلت الأوزان قبل وبعد أسبوعين

السيدة	1	2	3	4	5	6	7
الوزن قبل	128	132	136	150	140	137	124
الوزن بعد	129	120	127	135	128	130	159

أوجد 99% فترة ثقة لـــ μ_D إذا علم أن توزيعات الأوزان تقريباً طبيعية .

-٣٧- لمقارنة عليقتين من ناحية تأثيرهما على نمو العجول خلال شهرين من التغذيـــة اختـــيرت عينة عشوائية من 10 أزواج (توائم من العجول) لكل زوج أعطى إحدى التوأمين العليقـــة A و الآخر العليقة B وتم الحصول على النتائج التالية : $\bar{d} = 10$, $s_{\bar{d}} = 3$ أوجد 95% فـــترة ثقة للمعلمة μ_D.

-٣٨- طبق مقياس للاتجاهات نحو الأطفال في عينة من السيدات الآتي تم قبولهم بقسم دراسات الطفولة فور التحاقهن بالقسم ، ثم طبق نفس المقياس مرة أخرى فور حصولهن على البكالوريوس وتم الحصول على النتائج التالية $\bar{d} = 18$, $s_{\bar{d}} = 3$, $n = 10$ أوجد 99% فـــترة ثقـــة للمعلمة μ_D .

-٣٩- قام شخص بإجراء 10 عمليات حسابية على آلتين حاسبتين وتسجيل الزمن اللازم لكل عملية على كل آلة وقد تم الحصول على البيانات التالية $s_{\bar{d}} = 4$, $\bar{d} = 9$ أوجد 99% فترة ثقة للمعلمة μ_D .

٤٠- في عينة عشوائية من 1000 شخص من القراء لصحيفة ما وجد أن 400 منهم يفضلون قراءة نوع معين من الإعلانات في الصحيفة أوجد 99% فترة ثقة للنسبة p.

٤١- في مباراة رياضية للجرى وجد أن 240 طالب من 400 طالب يمكنهم الجري لمدة ميل في أقل من 7 دقائق . أوجد 95% فترة ثقة للنسبة p.

٤٢- أراد أحد مكاتب الاستقصاء معرفة نسبة الأصوات التي تؤيد مرشحاً معيناً في الانتخابات. اختيرت عينة عشوائية من 100 شخص فوجد أن من بينهم 55 فرداً يؤيدون هذا المرشح أوجد 99% فترة ثقة للنسبة p .

٤٣- أخذت عينة عشوائية من 200 طالب فكان عدد الناجحين منهم 144طالب . المطلوب

(أ) نسبة النجاح في العينة

(ب) 95% فترة ثقة لنسبة النجاح p .

٤٤- قدر حجم العينة العشوائية التي يمكن اختيارها لتقدير نسبة الطلاب الناجحين في مادة الإحصاء إذا كانت نسبة النجاح في عينة مبدئية 0.75 وبشرط أن أقصى خطأ لا يزيد عن 0.05 بثقة 95% .

٤٥- في إحدى مؤسسات رعاية الأحداث بها 6000 نزيل ، أخذت عينة عشوائية مكونة من 300 وأجريت دراسة لتحديد سبب دخول الحدث المؤسسة فوجد أن 80% من العينة يرجع سبب الدخول للمؤسسة إلى عدم رعاية الأم له. أوجد 99% فترة ثقة لنسبة الأحداث في المؤسسة الذين يرجع سبب دخولهم المؤسسة إلى عدم رعاية الأم.

٤٦- أخذت عينة عشوائية من 100 شخص من أعضاء هيئة التدريس ووجد أن 20% منهم أجورهم تزيد عن 30000 دولار في السنة. أوجد فترة ثقة للنسبة p .

٤٧- اتحاد جامعي يرغب في تقدير نسبة الطلبة الذين يفضلون هيئة طلابية جديدة. اختيرت عينة عشوائية من 400 طالب. فإذا أعطت نتائج التصويت $\hat{p} = 0.8$. أوجد 95% فترة ثقة للنسبة p .

٤٨- اختيرت عينة عشوائية من 400 مواطن في مجتمع سكاني ما ووجد أن 300 منهم يفضلون إضافة قليل من الفلور إلى مياههم. استخدم هذه البيانات في إيجاد 99% فترة ثقة للنسبة p.

٤٩-اختيرت عينة عشوائية من 600 مدخن سجائر في مجتمع سكاني ما ووجد أن 80 منهم يفضلون النوع A استخدم هذه البيانات في إيجاد :

(أ) نسبة النجاح في العينة

(ب) 95% فترة ثقة لنسبة النجاح p .

(ت) تقدير حجم العينة التي يمكن اختيارها لتقدير نسبة المدخنين الذين يفضلـــون النـــوع A وذلك بثقة 95% إذا كان حجم الخطأ 0.172.

-٥٠- في دراسة لنسبة ربات البيوت اللاتي يمتلكن غسالة بمجفف وجـــد أن 55 مـــن 100 سيدة في المدينة A يمتلكن غسالة بمجفف بينما في المدينة B وجد أن 45 مـــن 150 يمتلكـــن غسالة بمجفف. أوجد 95% فترة ثقة لـــ $p_1 - p_2$.

-٥١-أوجد 95% فترة ثقة للمعلمة σ^2 للمثال 19.

-٥٢-أوجد 95% فترة ثقة للمعلمة σ^2 للمثال 20.

-٥٣-أوجد 99% فترة ثقة للمعلمة $\frac{\sigma_1^2}{\sigma_2^2}$ للمثال 32.

-٥٤-أوجد 95% فترة ثقة للمعلمة $\frac{\sigma_1^2}{\sigma_2^2}$ للمثال 33.

الفصل التاسع

اختبارات الفروض

Tests of Hypotheses

(٩-١) الفروض الإحصائية Statistical Hypotheses

تعتبر اختبارات الفروض الإحصائية أهم فرع في نظرية القرارات. أولا، دعنا نعرف بدقة ماذا نعنى بالفرض الإحصائي .

تعريف : الفرض الإحصائي هو جملة ما تخص واحد أو أكثر من المجتمعات، من الممكن أن تكون صحيحه أو غير صحيحة .

للتأكد من صحة أو عدم صحة الفرض الإحصائي لا بد من دراسة كل مفردات المجتمع تحت الدراسة وهذا بالطبع غير عملي في معظم الحالات. بدلا من ذلك فإننا نختار عينة عشوائية من المجتمع ونستخدم المعلومات الموجودة في العينة لنتخذ قرار بقبول أو رفض الفرض الإحصائي. القرار الذي نتخذه سوف يكون سليم إذا كان الفرض صحيح وتم قبوله أو خطأ وتم رفضه. بينما يكون القرار غير سليم إذا كان الفرض صحيح وتم رفضه أو غير صحيح وتم قبوله.

الفروض التي نضعها على أمل أن نرفضها تسمى فروض العدم null hypotheses . ويرمز لفرض العدم بالرمز H_0 . رفض فرض العدم يؤدى إلى قبول فرض بديل hypothesis alternative و يرمز للفرض البديل بالرمز H_1 . فعلى سبيل المثال إذا كان فرض العدم H_0 أن متوسط الطول في مجتمع ما $\mu = 160$ (مقاسه بالسنتيمتر) فإن الفرض البديل H_1 قد يكون $\mu \neq 160$ أو $\mu < 160$ أو $\mu > 160$.

(٩-٢) الخطأ من النوع الأول والخطأ من النوع الثاني :

Type I Error and Type II Error

سوف نسهل المفاهيم المستخدمة في اختبارات الفروض والتي تخص مجتمع ما بالمثال التالي : بفرض أنه تم إجراء اختبار قدرات لعدد من المتقدمين لشغل وظيفة ما في مجال الحاسب الآلي. يشتمل الاختبار على 15سؤال وكل سؤال له 5 أجوبة ممكنة واحد منهم الصحيح. من شروط اجتياز الاختبار حصول المتقدم على 7 درجات فأكثر. هذا يعنى أن المتقدم لديه بعض المعلومات والتي تؤهله للعمل في مجال الحاسب الآلي. فرض العدم في هذه الحالة أن معلمة ذي الحدين (احتمال النجاح) في محاولة معطاة (سؤال) هو $p = \frac{1}{5}$ ، أى أن الشخص المتقدم للاختبار يعتمد على التخمين. الفرض البديل في هذه الحالة $p > \frac{1}{5}$. وعلى ذلك يمكن وضع فرض العدم والفرض البديل على الصورة التالية :

$$H_0 : p = \frac{1}{5}$$

$$H_1 : p > \frac{1}{5}$$

أن الطريقة السابقة في اتخاذ القرار قد تؤدى إلى استنتاجين غير صحيحين . فقد يحصل المتقدم للوظيفة على 7 درجات أو اكثر عن طريق التخمين. في هذه الحالة نكون قد وقعنا في خطأ عند رفض فرض العدم وقبول الفرض البديل مع أن H_0 صحيحا. مثل هذا الخطأ يسمى خطأ من النوع الأول.

تعريف : يحدث الخطأ من النوع الأول إذا كان فرض العدم صحيح ويتخذ قرار برفضه .

النوع الثاني من الخطأ والذي يمكن الوقوع فيه إذا حصل المتقدم للوظيفة على درجة أقل من 7 درجات وتستنتج أنه يخمن بينما هو في الحقيقة لديه بعض المعلومات عن الإجابة الصحيحة.

تعريف : يحدث الخطأ من النوع الثاني إذا قبلنا فرض العدم وهو خطأ.

يعتمد قرارنا في هذا المثال على الإحصاء X الذي يمثل عدد الإجابات الصحيحة التي يحصل عليها المتقدم في الاختبار حيث أن $x = 0, 1, 2, \ldots, 15$. القيم الممكنة من 0 إلى 15 تقسم إلى مجموعتين. تضم المجموعة الأولي القيم الأقل من 7 أما المجموعة الثانية فتضم القيم التي تساوي 7 فأكثر. كل الدرجات الممكنة والتي تساوى 7 فأكثر تكون منطقة الرفض region rejection بينما الدرجات التي أقل من 7 تكون منطقة القبول acceptance region . الرقم 7 يسمى القيمة الحرجة critical region. إذا وقعت قيمة الإحصاء الذي يعتمد عليه قرارنا في منطقة الرفض نرفض فرض العدم H_0 ونقبل الفرض البديل H_1 بينما إذا وقعت قيمة الإحصاء في منطقة القبول نقبل H_0 ونرفض H_1 .

تعريف : احتمال الوقوع في خطأ من النوع الأول يسمى مستوى المعنوية level of significance للاختبار ويرمز له بالرمز α .

في مثالنا فإن الخطأ من النوع الأول يقع إذا حصل المتقدم للاختبار على 7 درجات أو أكثر عن طريق التخمين. فإذا كانت X تمثل عدد الإجابات الصحيحة فإن :

$$\alpha = P\,(\text{الخطأ من النوع الأول})$$

$$= P\,(\text{رفض } H_0 \text{ وهو صحيح})$$

$$= P(X \geq 7 \mid p = \frac{1}{5})$$

$$= \sum_{x=7}^{15} b\left(x;\ 15,\ \frac{1}{5}\right)$$

$$= 1 - \sum_{x=0}^{6} b\left(x;\ 15,\ \frac{1}{5}\right)$$

$$= 1 - 0.982 = 0.018.$$

وهذا يعنى أنه تقريباً في 1.8% من كل التجارب من هذا النوع حيث $n = 15$ تؤدى إلى الوقوع في خطأ من النوع الأول، أى رفض H_0 وهو صحيحا. في هذه الحالة يمكن القول أن $H_0 : p = \frac{1}{5}$ تختبر عند مستوى معنوية $\alpha = 0.018$. أحيانا تسمى α حجم منطقة الرفض size of rejection .

احتمال الوقوع في خطأ من النوع الثاني يرمز له بالرمز β ولحساب قيمة β لابد من وضع فرض بديل معين. فمثلا عند اختبار فرض العدم $p = \frac{1}{5}$ ضد الفرض البديل $p = \frac{7}{10}$ فإننا نكون قادرين على حساب احتمال حصول المتقدم للاختبار على درجة أقل من 7 عندما تكون $p = \frac{7}{10}$ ، في هذه الحالة :

$$\beta = P\ (\text{خطأ من النوع الثاني})$$

$$= P\ (\text{قبول } H_0 \text{ وهو خطأ})$$

$$= P\left(X < 7 \mid p = \frac{7}{10}\right)$$

$$= \sum_{x=0}^{6} b\left(x;\ 15,\ \frac{7}{10}\right) = 0.015$$

وهذا يعنى أن تقريباً 1.5% من كل التجارب من هذا النوع حيث $n = 15$ تؤدى إلى قبول فرض العدم وهو خطأ. في الحقيقة لكل قيمة من α لا يوجد قيمة وحيدة من β بل يوجد قيم مختلفة لـ β وذلك لقيمة $p = \frac{1}{5}$. على سبيل المثال يوجد قيمة لـ β عند $p = 0.6$ وقيمة لـ β عند $p = 0.7$ وهكذا. الجدول التالي يعطى قيم β لقيم مختلفة مختارة من p (كل واحدة تمثل الفرض البديل لفرض العدم $H_0 : p = \frac{1}{5}$).

p	0.6	0.7	0.8	0.9
β	0.095	0.015	0.001	0.000

يُلاحظ من الجدول أنه كلما بعدت قيمة p (تحت الفرض البديل) عـــن $p = \frac{1}{5}$ كلمـــا قـــل الوقوع في خطأ من النوع الثاني أي كلما قلت قيمة β عند منطقة الرفض $X \geq 7$.

مثال (٩-١) إذا كانت p تمثل نسبة الأصوات المؤيدة للشخص A ضد الشخص B في ترشيح ما. اختيرت عينة عشوائية من الحجم $n = 20$ فإذا كانت X تمثل عـــدد الأشـــخاص الذين يؤيدون الشخص A في العينة .

المطلوب : (أ) حساب احتمال الوقوع في خطأ من النوع الأول إذا كان $H_0: p = 0.5$ ضـــد الفرض البديل $H_0 : p \neq 0.5$ ومنطقة الرفض $X \leq 5$ أو $X \geq 12$.

(ب) حساب احتمال الوقوع في الخطأ من النـــوع الثـــاني تحـــت الفـــرض البديـــل $H_1: p = 0.9$.

الحل . (أ) $\alpha = P$ (الخطأ من النوع الأول)

$= P$ (رفض H_0 وهو صحيح)

$$= P(X \geq 12 \mid p = 0.5) + P(X \leq 5 \mid p = 0.5)$$

$$= (1 - \sum_{x=0}^{11} b\ (x;\ 20, 0.5)) + \sum_{x=0}^{5} b\ (x;\ 20, 0.5)$$

$$= (1 - 0.748) + 0.021$$

$$= 0.252 + 0.021 = 0.273$$

(ب) $\beta = P$ (الخطأ من النوع الثاني)

$= P$ (قبول H_0 وهو خطأ)

$$= P(5 < X < 12 \mid p = 0.9) = p(6 \leq X \leq 11 \mid p = 0.9)$$

$$= \sum_{x=0}^{11} b\ (x;\ 20, 0.9) - \sum_{x=0}^{5} b\ (x;\ 20, 0.9)$$

$$= 0.0000 - 0.0000 = 0.0000 .$$

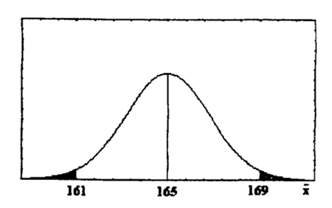
عموماً الاختبار الجيد هو الذي يجعل كلا من α, β أصغر ما يمكن ومن الصعب الحصول علـــى هذا الاختبار لأنه لا يمكن تصغير كل من α, β في آن واحد حيث أن تصغير إحداهما يـــؤدى إلى تكبير الأخرى. لذلك لجأ الإحصائيون إلى تثبيت مستوى المعنوية α عند قيمة محددة ثم اختيـــار الاختبار الإحصائي الذي يجعل β أقل ما يمكن. من القيم الشائعة لمستوى المعنوية $\alpha = 0.05$ أو $\alpha = 0.01$. يقال للاختبار أنه معنوي significant عند $\alpha = 0.05$ إذا رفض فرض العدم

عند مستوى معنوية $\alpha = 0.05$ ويعتبر الاختبار معنوي جدا إذا رفض فرض العدم عند مستوي معنوية $\alpha = 0.01$.

من السهل توضيح المفاهيم السابقة بيانياً عندما يكون المجتمع متصلا. على سبيل المثال إذا كان فرض العدم أن متوسط الطول لمجموعة من الطلبة في جامعة ما هو $\mu = 165$ سم ضد الفرض البديل أن متوسط الطول لا يساوى 165سم . أى أننا نرغب في اختبار :

$$H_0 : \mu = 165,$$
$$H_1 : \mu \neq 165.$$

الفرض البديل يمكن أن يكون $\mu < 165$ أو $\mu > 165$. بفرض أن الانحراف المعياري لمجتمع الأطوال $\sigma = 10$ ، الإحصاء الذي سوف نبنى عليه قرارنا والذي يعتمد على عينة عشوائية من الحجم $n = 30$ سوف يكون $\overline{X}$ والذي يعتبر التقدير الأكثر كفاءة للمعلمة μ . علمنا في الفصل السابع أن التوزيع العيني للإحصاء $\overline{X}$ تقريباً يتبع التوزيع الطبيعي بانحراف معياري $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 10/\sqrt{30} = 1.8$. بفرض أننا اخترنا منطقة الرفض ممثلة في قيم $\bar{x}$ الأقل من 161 أو قيم $\bar{x}$ التي أكبر من 169. أى أن $\overline{X} < 161$ أو $\overline{X} > 169$ والموضحة في شكل (٩-١) . منطقة القبول سوف تكون $161 < \overline{X} < 169$. وعلى ذلك إذا وقع متوسط العينة $\bar{x}$ في منطقة القبول نقبل H_0 ونرفض H_1 وغير ذلك نرفض H_0 . احتمال الوقوع في خطأ من النوع الأول، أو مستوى المعنوية للاختبار، يساوى مجموع المساحات المظللة في كل من جانبي التوزيع العيني للإحصاء $\overline{X}$ في شكل (٩-١).



شكل (٩ – ١)

وعلى ذلك فإن :

$$\alpha = P(\ \overline{X} < 161 \mid \text{صحيح } H_0) + P(\ \overline{X} > 169 \mid \text{صحيح } H_0)$$

القيمتين الحرجتين للمتغير العشوائي Z والمقابلتين للقيمتين $\bar{x}_1 = 161$ ، $\bar{x}_2 = 169$, عندما تكون H_0 صحيحةً هما :

$$z_1 = \frac{161 - 165}{1.8} = -2.22,$$

$$z_2 = \frac{169 - 165}{1.8} = 2.22.$$

وعلى ذلك فإن :

$$\begin{aligned}\alpha &= P(Z < -2.22) + P(Z > 2.22) \\ &= 2P(Z > 2.22) \\ &= 2[0.5 - P(0 < Z < 2.22)] \\ &= 2[0.5 - 0.4868] = 2(0.0132) \\ &= 0.0264 .\end{aligned}$$

مثال (٩-٢) إذا كان الزمن اللازم لجفاف طلاء من نوع ما يتبع توزيعاً طبيعياً بمتوسط 75 دقيقه وانحراف معياري $\sigma = 9$ دقيقة. فإذا أضيفت تعديلات على هذا الطلاء وذلك لتقليل زمن الجفاف وإذا كان الطلاء بعد التعديل لا يزال يتبع توزيعاً طبيعياً بنفس الانحراف المعياري (σ=9) وإذا كان μ تمثل متوسط زمن الجفاف للطلاء بعد التعديل .

المطلوب : (أ) حساب احتمال الوقوع في خطأ من النوع الأول إذا كان $H_0 : \mu = 75$ و $H_1 : \mu < 75$ وذلك بالاعتماد على عينة عشوائية حجمها $n = 25$ ومنطقة الرفض $\overline{X} \le 70.8$.

(ب) حساب احتمال الوقوع في خطأ من النوع الثاني تحت فرض العدم $H_0: \mu$ =75 ضد الفرض البديل أن 72 و $H_1: \mu = 70$.

الحل . (أ) $\alpha = P$ (الخطأ من النوع الأول)

$= P$ (رفض H_0 وهو صحيح)

$= P(\overline{X} \le 70.8 \mid \mu = 75)$

قيمة z المقابلة لقيمة $\bar{x} = 70.8$ في هذه الحالة هي :

$$z = \frac{70.8 - 75}{9/\sqrt{25}} = -2.33,$$

وعلى ذلك يكون :

$$\begin{aligned}\alpha &= P(Z < -2.33) = P(Z > 2.33) \\ &= 0.5 - P(0 < Z < 2.33)\end{aligned}$$

$= 0.5 - 0.4901 = 0.0099.$

(ب) حساب β عندما $H_1: \mu = 72$

$\beta = P$ (الخطأ من النوع الثاني)

$= P$ (قبول فرض العدم وهو خطأ)

$= P (\overline{X} > 70.8 \mid \mu = 72)$

قيمة z المقابلة لقيمة $\overline{x} = 70.8$ في هذه الحالة هي :

$$z = \frac{70.8 - 72}{9/\sqrt{25}} = -0.67,$$

وعلى ذلك تكون :

$$\beta = P (Z > -0.67)$$
$$= 0.5 + P (0 < Z < 0.67)$$
$$= 0.5 + 0.2486 = 0.7486.$$

حساب β عند $H_1: \mu = 70$

$\beta = P$ (الخطأ من النوع الثاني)

$= P$ (قبول فرض العدم وهو خطأ)

$= P (\overline{X} > 70.8 \mid \mu = 70)$

قيمة z المقابلة لقيمة $\overline{x} = 70.8$ في هذه الحالة هي :

$$z = \frac{70.8 - 70}{9/\sqrt{25}} = 0.44,$$

وعلى ذلك تكون :

$$\beta = P (Z > 0.44)$$
$$= 0.5 - P (0 < Z < 0.44)$$
$$= 0.5 - 0.17 = 0.33.$$

(٩-٣) اختبارات من جانب واحد أو من جانبين :

One – tailed and Two – tailed Tests

يسمى الاختبار ، لأى فرض إحصائي ، أنه من جانب واحد إذا كان على الصورة :

$$H_0 : \theta = \theta_0 ,$$
$$H_1 : \theta > \theta_0$$

أو على الشكل :

$$H_0 : \theta = \theta_0 ,$$
$$H_1 : \theta < \theta_0$$

منطقة الرفض للبديل $\theta > \theta_0$ تقع في الجانب الأيمن من التوزيع بينما منطقة الرفض للبديل $\theta < \theta_0$ تقع في الجانب الأيسر من التوزيع.

يسمى الاختبار، لأي فرض إحصائي، أنه من جانبين إذا كان على الصورة :

$$H_0 : \theta = \theta_0 ,$$
$$H_1 : \theta \neq \theta_0$$

منطقة الرفض للبديل $\theta \neq \theta_0$ سوف تكون $\theta_0 > \theta$ أو $\theta < \theta_0$. المفاضلة بين اختبار من جانب واحد أو من جانبين سوف يتوقف على الاستنتاج الذي يرغب الباحث في الوصول إليه عند رفض فرض العدم :

في البنود التالية من هذا الفصل سوف نناقش بعض اختبارات الفروض الشائعة الاستخدام.

(٩-٤) اختبارات حول متوسط المجتمع μ

Tests About a Population Mean μ

الحالة الأولى : اختبار الفرض أن المتوسط لمجتمع بتباين معلوم σ^2 ، يساوى قيمة معينة μ_0 ضد الفرض البديل ذي جانبين أن المتوسط لا يساوى μ_0 يكون على الشكل:

$$H_0 : \mu = \mu_0 ,$$
$$H_1 : \mu \neq \mu_0$$

الإحصاء المناسب والذي يعتمد عليه قرارانا هو المتغير العشوائي $\overline{X}$. إذا كان المجتمع الذي اختيرت منه العينة يتبع التوزيع الطبيعي فإن التوزيع العيني للإحصاء $\overline{X}$ يتبع التوزيع الطبيعي بمتوسط $\mu_{\bar{X}} = \mu$ وتباين $\sigma^2_{\bar{X}} = \dfrac{\sigma^2}{n}$ حيث أن μ و σ^2 هما المتوسط والتباين على التوالي للمجتمع الذي اختيرت منه العينات من الحجم n . عندما لا يتحقق هذا الفرض فقد علمنا من الفصل السابع أن التوزيع العيني للإحصاء $\overline{X}$ تقريباً يتبع التوزيع الطبيعي بمتوسط $\mu_{\bar{X}} = \mu$ وتباين $\sigma^2_{\bar{X}} = \dfrac{\sigma^2}{n}$. باستخدام القيمتين الحرجتين $\bar{x}_1$, $\bar{x}_2$ فإن $\bar{x}_1 < \overline{X} < \bar{x}_2$ تمثل منطقة القبول. الذيلين للتوزيع $\overline{X} > \bar{x}_2$ أو $\overline{X} < \bar{x}_1$ يمثلان منطقة الرفض .

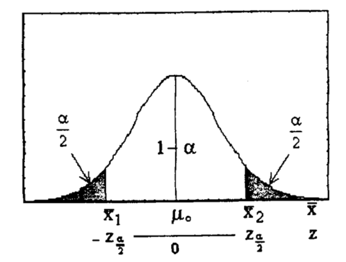
منطقة الرفض يمكن أن تُعطى في صورة قيم z وذلك بعمل التحويلة التالية :

$$z = \frac{\bar{x} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}} .$$

حيث ، لمستوى معنوية α ، القيمتين الحرجتين للمتغير العشوائي Z المقابلة لكل من $\bar{x}_1$, $\bar{x}_2$ والموضحَين في شكل (٩-٢) هما :

$$-z_{\alpha/2} = \frac{\bar{x}_1 - \mu_0}{\frac{\sigma}{\sqrt{n}}},$$

$$z_{\alpha/2} = \frac{\bar{x}_2 - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$



شكل (٩ - ٢)

الآن لإجراء الاختبار نختار عينة عشوائية من الحجم n من المجتمـــع موضـــع الدراســـة ونحسب منها متوسط العينة $\bar{x}$. إذا وقعت $\bar{x}$ في منطقة القبـــــول $\bar{x}_1 < \bar{X} < \bar{x}_2$ فإن قيمة الإحصاء :

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

تقع في المنطقة $-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}$. الاستنتاج سوف يكون $\mu = \mu_0$ وغير ذلك نرفـض H_0 ونقبل الفرض البديل $\mu \neq \mu_0$. عادةً تصا غ منطقة الرفض في صورة Z أكثر من $\bar{X}$. إذا كان تباين المجتمع مجهول فإننا نحسب تباين العينةs^2 ونستخدمه بدلاً من σ^2 تحت شـــرط أن حجـــم العينة أكبر من أو يساوى 30 ($n \geq 30$) .

مثال (٩-٣) ينتج مصنع للأغذية المعلبة نوعا من المعلبات. قام المسئولين خلال فترة طويلة بمراقبة أوزان هذه المعلبات ووجد أنها تخضع للتوزيع الطبيعي بانحراف معياري 2.6 جرام. جرت العادة في المصنع أن يكتب على العلبة الوزن الصافي وهو 300 جرام. اختيرت عينة عشوائية من 20 علبة وكان متوسط الوزن من العينة $\bar{x} = 305$. أختبر فرض العدم $\mu = 300$ ضد الفرض البديل $\mu \neq 300$ وذلك عند مستوى معنوية $\alpha = 0.01$.

الحل .

$$H_0 : \mu = 300,$$
$$H_1 : \mu \neq 300.$$
$$\alpha = 0.01 .$$

$z_{0.005} = 2.575$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣)

منطقة الرفض $Z > 2.575$ أو $Z < -2.575$

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{305 - 300}{\frac{2.6}{\sqrt{20}}} = 8.6.$$

نرفض فرض العدم لأن z تقع في منطقة الرفض.

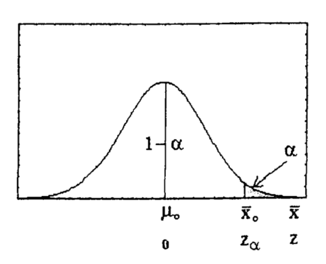
عند الاهتمام باختبار الفرض أن متوسط المجتمع بتباين معلوم ، σ^2 ، يساوى قيمة معينة μ_0 ضد الفرض البديل من جانب واحد أي $\mu > \mu_0$ فإن فرض العدم والفرض البديل سوف يكونان على الشكل :

$$H_0 : \mu = \mu_0,$$
$$H_1 : \mu > \mu_0.$$

منطقة الرفض سوف تقع في الجانب الأيمن من توزيع $\bar{X}$. لمستوى معنوية α نحسب قيمة حرجة واحدة $\bar{x}_0$ بحيث $\bar{X} < \bar{x}_0$ تمثل منطقة القبول بينما $\bar{X} > \bar{x}_0$ تمثل منطقة الرفض . في شكل (٩-٣). القيمة الحرجة z_α التي تقابل القيمة $\bar{x}_0$ هى :

$$z_\alpha = \frac{\bar{x}_0 - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

وعلى ذلك منطقة الرفض من الحجم α تصبح $Z > z_\alpha$. للفرض البديل $\mu < \mu_0$ سوف نتبع خطوات الطريقة السابقة فيما عدا أن منطقة الرفض سوف تكون $Z < -z_\alpha$.



شكل (٩-٣)

مثال (٩-٤) من المعروف أن أحد أدوية إزالة الألم المستخدمة يمكنها إزالة الألم للمريض في فترة زمنية متوسطها 3.7 دقيقة . ولمقارنة هذا الدواء بدواء جديد لإزالة الألم اختيرت عينة عشوائية من 60 مريضا وتم إعطاء الدواء الجديد لهم فكان المتوسط الحسابي لطول فترة إزالة الألم في هذه العينة 2.2 دقيقة بانحراف معياري 1.2 دقيقة . فهل تدل هذه النتائج أن الدواء الجديد أفضل من الدواء القديم من حيث الفترة اللازمة لإزالة المرض ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل .

$$H_0 : \mu = 3.7 ,$$
$$H_1 : \mu < 3.7 .$$
$$\alpha = 0.05 .$$

وبما أن $n \geq 30$ فإنه يمكننا استخدام s بدلا من σ في صيغة z .

أن $z_{0.05} = 1.645$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣)

منطقة الرفض $Z < -1.645$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{2.2 - 3.7}{\frac{1.2}{\sqrt{60}}} = -9.682.$$

نرفض فرض العدم لأن z تقع في منطقة الرفض. أى أن الدواء الجديد أفضل من الدواء القديم من حيث الفترة اللازمة لإزالة المرض.

مثال (٩-٥) اتفق أحد مصدري البيض مع أحد التجار على أن يورد الأخير للأول عدد من البيض من الحجم الكبير ولما أحضر التاجر البيض قام المصدر باختيار عينة عشوائية من 100 بيضة فوجد أن متوسط وزن البيضة 67 جراما بانحراف معياري 1.6. اختبر فرض العدم $H_0 : \mu = 65$ ضد الفرض البديل $H_1 : \mu > 65$ (مستوى المعنوية $\alpha = 0.05$)

الحل .

$$H_0 : \mu = 65,$$
$$H_1 : \mu > 65.$$
$$\alpha = 0.05.$$

بما أن $n > 30$ فإنه يمكننا استخدام s بدلا من σ في صيغة z.

الحل .

$z_{0.05} = 1.645$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) .

منطقة الرفض $Z > 1.645$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{67 - 65}{\frac{1.6}{\sqrt{100}}} = 12.5.$$

نرفض H_0 لأن z تقع في منطقة الرفض. أي أن المورد سوف يقبل استلام البيض .

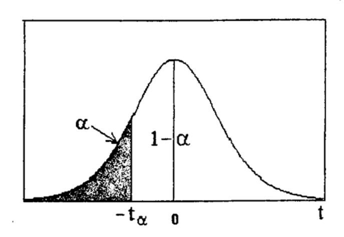
الحالة الثانية : في بعض اختبارات الفروض التي تخص متوسط مجتمع طبيعي عندما يكون تباين المجتمع σ^2 مجهول وحجم العينة صغير $n < 30$. استنتاجنا، في هذه الحالة، سوف يعتمد على توزيع t . بفرض اختبار من جانب واحد على الشكل :

$$H_0 : \mu = \mu_0,$$
$$H_1 : \mu < \mu_0$$

نختار عينة عشوائية من الحجم $n < 30$ من المجتمع ونحسب المتوسط $\bar{x}$ والانحراف المعياري s وعلى ذلك :

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائي T له توزيع t بدرجات حرية $v = n - 1$ عندما H_0 صحيحاً. منطقة الرفض سوف تكون في الذيل الأيسر من توزيع t . لمستوى معنوية α يمكن الحصول على قيمة واحدة $t_{1-\alpha} = -t_\alpha$ بحيث أن $T < -t_\alpha$ تمثل منطقة الرفض و $T > -t_\alpha$ تمثل منطقة القبول. حجم منطقة الرفض يساوى المساحة المظللة في شكل (٩-٤) .



سكل (٩-٤)

منطقة الرفض للفرض البديل $H_0 : \mu > \mu_0$ بمستوى معنوية α ، هى $T > t_\alpha$. للفرض البديل $H_0 : \mu \neq \mu_0$ فإن منطقة الرفض المقابلة لمستوى معنوية α ، هى $T > t_{\alpha/2}$ أو $T < -t_{\alpha/2}$. وعلى ذلك نحسب قيمة الإحصاء ، أي قيمة t ، وإذا وقعت قيمة t في منطقة القبول نقبل فرض العدم وغير ذلك نرفض H_0 .

مثال (٩-٦) لمعرفة اثر غذاء معين على زيادة الوزن اختيرت عينة عشوائية من ستة فئران وتم تغذيتها بهذا الغذاء وكانت الزيادة في أوزانهم بعد التغذية هى :

$$2.3 \ , 2.5 \ , 1.4 \ , 1.4 \ , 1.7 \ , 2.5$$

فهل يمكن الحكم على أن هذه العينة من مجتمع متوسط الزيادة في الوزن فيه 1.2 أم لا ؟ وذلك عند مستوى معنوية $\alpha = 0.05$. وذلك تحت فرض أن العينة تم اختيارها من مجتمع طبيعي .

الحل .

$$H_0 : \mu = 1.2 \ ,$$
$$H_0 : \mu \neq 1.2 \ .$$
$$\alpha = 0.05 \ .$$

$t_{.025} = 2.571$ والمستخرجة من توزيع t في ملحق (٤) عند درجات حرية $v = 5$.

منطقة الرفض $T > 2.571$ أو $T < -2.571$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{11.8}{6} = 1.967,$$

$$s = \sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}\right]}$$

$$= \sqrt{\frac{1}{5}\left[24.6 - \frac{(11.8)^2}{6}\right]} = 0.528.$$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{1.967 - 1.2}{\frac{0.528}{\sqrt{6}}} = 3.5582.$$

نرفض H_0 لأن t تقع في منطقة الرفض .

(٩-٥) اختبارات حول تباين المجتمع σ^2

Tests about the Population Variance σ^2

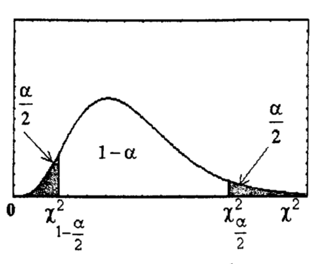
عند الرغبة في اختبار الفرض أن التباين لمجتمع طبيعي يساوى قيمة معينة σ_0^2 ضد الفرض البديل ذي جانبين أن التباين لا يساوى σ_0^2 . أي أننا نختبر الفرض أن :

$$H_0: \sigma^2 = \sigma_0^2 ,$$
$$H_1: \sigma^2 \neq \sigma_0^2$$

نختار عينة عشوائية من الحجم n من المجتمع موضع الدراسة ونحسب تباين العينة s^2. وعلى ذلك:

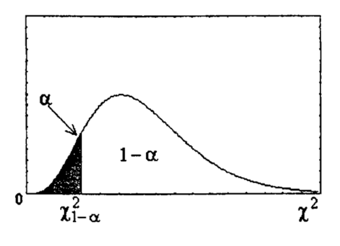
$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

قيمة للمتغير X^2 ، تبعا لنظرية (٧ – ١٢)، والذي له توزيع χ^2 بدرجات حرية $v = n - 1$ عندما يكون H_0 صحيحاً. لمستوى معنوية α نوجد القيمتين الحرجتين $\chi^2_{1-\frac{\alpha}{2}}, \chi^2_{\frac{\alpha}{2}}$ بحيث أن $X^2 < \chi^2_{1-\frac{\alpha}{2}}, X^2 > \chi_{\frac{\alpha}{2}}$ يمثلان منطقة الرفض. حجم المنطقة الحرجة يساوى المساحة المظللة في شكل (٩-٥). نرفض H_0 إذا وقعت χ^2 في منطقة الرفض .



شكل (٩-٥)

عادة يكون الاهتمام باختبار الفرض $\sigma^2 = \sigma_0^2$ ضد الفرض البديل من جانب واحد من التوزيع . للبديل $\sigma^2 < \sigma_0^2$ ولمستوى معنوية α نحصل على قيمة حرجة $\chi^2_{1-\alpha}$ بحيث أن $X^2 < \chi^2_{1-\alpha}$ تمثل منطقة الرفض و $X^2 > \chi^2_{1-\alpha}$ تمثل منطقة القبول. بنفس الشكل لبديل من جانب واحد $\sigma^2 > \sigma_0^2$، فإن χ^2_{α} تمثل القيمة الحرجة بحيث $X^2 > \chi^2_{\alpha}$ تمثل منطقة الرفض و $X^2 < \chi^2_{\alpha}$ تمثل منطقة القبول. حجم المنطقة الحرجة لبديل من جانب واحد $\sigma^2 < \sigma_0^2$ يساوى المساحة المظللة في شكل (٩-٦) .



شكل (٩-٦)

مثال (٩-٧) يعتقد مسئول في مصنع لبطاريات السيارات أن الانحراف المعياري لعمر البطارية المنتجة هو 0.8 . لاختبار ذلك الفرض اختيرت عينة عشوائية من الحجم $n = 20$ وتمت تجربتها فكان $s = 1.1$ سنة فهل يمكن القول أن $\sigma > 0.8$ ؟(استخدم مستوى معنوية $\alpha = 0.05$).

الحل .

$$H_0: \sigma^2 = 0.64,$$
$$H_1: \sigma^2 > 0.64$$
$$\alpha = 0.05.$$

$\chi^2_{.05} = 30.143$ والمستخرجة من جدول توزيع χ^2 بدرجات حرية $v = 19$. منطقة الرفض $X^2 > 30.143$.

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(19)(1.1)^2}{0.8^2}$$
$$= 35.922.$$

بما أن χ^2 تقع في منطقة الرفض نرفض H_0.

(٩-٦) اختبارات تخص تبايني مجتمعين

Tests Concerning Two Populations Variances

بفرض أن لدينا مجتمعين : الأول يتبع توزيعاً طبيعياً متوسطة μ_1 وتباينــه σ_1^2 والثـاني : يتبع توزيعاً طبيعياً متوسطة μ_2 وتباينه σ_2^2 والمطلوب اختبار هل المجتمعين لهما نفس التباين ؟ أى هل $\sigma_1^2 = \sigma_2^2$ أم لا ؟ فإذا كانت $\sigma_1^2 = \sigma_2^2$ فإننا نقول أن هناك تجانس بين المجتمعين . إن التأكد من صحة الفرض $\sigma_1^2 = \sigma_2^2$ ضروري لاختبار الفرق بين متوسطي مجتمعين (اختبار t) والذي سوف نتناوله في البند التالي . أيضا هناك العديد من الأبحاث التي يكون هدفها الرئيسـي هو مقارنة σ_1^2 مع σ_2^2 مثل دراسات جودة البضائع المستهلكة حيث يعتـبر التبـاين أهـم مقاييس الجودة.

لاختبار فرض العدم :

$$H_0: \ \sigma_1^2 = \sigma_2^2$$

ضد الفرض البديل :

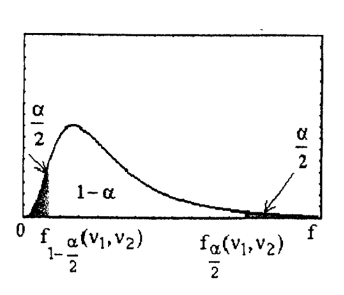
$$H_1: \ \sigma_1^2 \neq \sigma_2^2$$

نختار عينة عشوائية حجمها n_1 من المجتمع الأول وليكن متوسطها الحسابي $\bar{x}_1$ وتباينها s_1^2 وتختار عينة عشوائية أخرى حجمها n_2 من المجتمع الثاني وليكن متوسطها $\bar{x}_2$ وتباينها s_2^2 .
(العينة الثانية مستقلة عن العينة الأولى) . بافتراض صحة فرض العدم فإن :

$$f = \frac{s_1^2}{s_2^2}$$

تمثل قيمة للمتغير العشوائي F والذي تبعا لنظرية (٧-١٣) له توزيـــع F بدرجـات حرية $v_1 = n_1 - 1, v_2 = n_2 - 1$ • لمستوى معنوية α، سوف نحصـــل علـــى قيمتـــين حرجتين $f_{\frac{\alpha}{2}}(v_1, v_2)$ و $f_{1-\frac{\alpha}{2}}(v_1, v_2)$. وعلـــى ذلـــك فـــإن $F > f_{\frac{\alpha}{2}}(v_1, v_2)$ أو $F < f_{1-\frac{\alpha}{2}}(v_1, v_2)$ تمثلان منطقة الرفض. حجم منطقة الرفض يساوى المساحة المظللـــة في شكل (٩-٧). باستخدام نظرية (٧-١٤) فإن القيمة الحرجة للمتغيرF في الطرف الأيســـر يمكن الحصول عليها من العلاقة التالية :

$$f_{1-\frac{\alpha}{2}}(v_1, v_2) = \frac{1}{f_{\frac{\alpha}{2}}(v_2, v_1)} .$$



شكل (٩-٧)

إذا وقعت f المحسوبة في منطقة الرفض فإننا نرفض H_0 .

لاختبار فرض العدم :

$$H_0 : \ \sigma_1^2 = \sigma_2^2$$

ضد الفرض البديل :

$$H_1 : \ \sigma_1^2 < \sigma_2^2$$

فإن منطقة الرفض ، بمستوى معنوية α ، سوف تكون في الجانب الأيسر من التوزيع (الذيـــل الأيسر) . منطقة الرفض في هذه الحالة تمثل كل قيم F بحيث $F < f_{1-\alpha}(\nu_1, \nu_2)$. وأخيرا لاختبار فرض العدم :

$$H_0 : \sigma_1^2 = \sigma_2^2$$

ضد الفرض البديل :

$$H_1 : \sigma_1^2 > \sigma_2^2$$

فإن منطقة الرفض ، بمستوى معنوية α ، سوف تكون في الجانب الأيمن من التوزيـــع (الذيـــل الأيمن) . منطقة الرفض في هذه الحالة تمثل كل قيم F بحيث $F > f_{\alpha}(\nu_1, \nu_2)$.

مثال (٩-٨) من البيانات التالية اختبر التجانس بين المجتمعين وذلك عند مســـــتوى معنويـــة $\alpha = 0.1$

العينة الأولي	العينة الثانية	
40.5	50.7	s_i^2
31	41	n_i

الحل .

$$H_0 : \sigma_1^2 = \sigma_2^2 ,$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2 .$$
$$\alpha = 0.1 .$$

$f_{.05}(40, 30) = 1.79$ والمستخرجة من جدول توزيع F في ملحـــــق (٦) عنـــد درجات حرية $\nu_1 = 40, \nu_2 = 30$. أما $f_{0.95}(40, 30)$ فتحسب من العلاقة التالية :

$$f_{0.95}(40,30) = \frac{1}{f_{0.05}(30,40)} = \frac{1}{1.74} = 0.575.$$

منطقة الرفض $F > 1.79$ أو $F < 0.575$

$$f = \frac{s_2^2}{s_1^2} = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}} = \frac{50.7}{40.5} = 1.252.$$

نقبل H_0 لأن f تقع في منطقة القبول .

(٩-٧) اختبارات تخص المتوسطات Tests Concerning Means

في بعض الأحيان يكون الاهتمام باختبارات الفروض التي تخص مجتمعين مختلفين. أي أننا نرغب في اختبار فرض العدم أن الفرق بين متوسطي مجتمعين ، $\mu_1 - \mu_2$ ، يساوى صفر أي $\mu_1 = \mu_2$ ضد الفرض البديل $\mu_1 - \mu_2 > 0$ أي $\mu_1 > \mu_2$ أو الفرض البديل $\mu_1 - \mu_2 < 0$ أي $\mu_1 < \mu_2$ أو الفرض البديل $\mu_1 - \mu_2 \neq 0$ أي $\mu_1 \neq \mu_2$. تعتمد الطريقة المستخدمة في اختيار الفرق بين متوسطي مجتمعين على توزيع كل مجتمع وحجم العينة المختارة من كل مجتمع. في الجزء التالي سوف نتناول ثلاثة حالات.

الحالة الأولي : عند اختبار فرض العدم H_0 أن الفرق بين متوسطي مجتمعين ، $\mu_1 - \mu_2$ ، يساوى صفر وذلك عندما كل من σ_1^2 , σ_2^2 معلومتان وتحت فرض أن كل مجتمع له توزيعاً طبيعياً أو تقريبا طبيعياً. أما في حالة العينات الكبيرة وإذا كانت σ_1^2 , σ_2^2 مجهولتان فإنه يمكن تقديرهما من العينات بحساب s_1^2 , s_2^2. يعتمد قرارنا في هذه الحالة على المتغير العشوائي (الإحصاء) $\overline{X}_1 - \overline{X}_2$. أولا نختار عينة عشوائية من الحجم n_1 من المجتمع الأول ونحسب منها $\overline{x}_1$ ونختار عينة عشوائية أخرى من الحجم n_2 من المجتمع الثاني (مستقلة عن العينة الأولى) ونحسب منها $\overline{x}_2$ ثم نحسب الفرق ، $\overline{x}_1 - \overline{x}_2$ ، لمتوسطي العينتين ، من نظرية (٧-٦) نعلم أن :

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - 0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}.$$

قيمة للمتغير العشوائي Z عندما يكون H_0 صحيحا. وعلى ذلك في اختبار من جانبين وعند مستوى معنوية α فإن منطقة الرفض تحدد على الشكل $Z > z_{\frac{\alpha}{2}}$ أو $Z < -z_{\frac{\alpha}{2}}$. أما في اختبار من جانب واحد حيث الفرض البديل $\mu_1 < \mu_2$ فإن منطقة الرفض ، لمستوى معنوية α ، سوف تكون $Z < -z_\alpha$. وأخيرا في حالة الفرض البديل من جانب واحد $\mu_1 > \mu_2$ فإن منطقة الرفض لمستوى معنوية α ، سوف تكون $Z > z_\alpha$.

مثال (٩-٩) أجرى اختبار على المقاومة للشد **tensile strength** لنوعين من السلك . النتائج معطاة في الجدول التالي :

النوع	حجم العينة	متوسط العينة	الانحراف المعياري للعينة
A	$n_1 = 129$	$\bar{x}_1 = 107.6$	$s_1 = 1.3$
B	$n_2 = 129$	$\bar{x}_2 = 123.6$	$s_2 = 2.0$

المطلوب اختبار هل هناك فرقا معنويا بين متوسطي المجتمعين المسحوبتين منهما العينتين ؟ (عند مستوى معنوية $\alpha = 0.05$).

الحل . حيث أن $n_1 > 30$ و $n_2 > 30$ نتبع الآتي :

$$H_0 : \mu_1 = \mu_2 ,$$

$$H_1 : \mu_1 \neq \mu_2 .$$

$$\alpha = 0.05 .$$

$z_{0.025} = 1.96$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) .

منطقة الرفض $Z > 1.96$ أو $Z < -1.96$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

$$= \frac{107.6 - 123.6}{\sqrt{\frac{1.3^2}{129} + \frac{2^2}{129}}} = -76.183.$$

وبما أن z تقع في منطقة الرفض فإننا نرفض H_0 .

الحالة الثانية : بفرض أن σ_1^2 , σ_2^2 مجهولتان وحجم كلا من العينتين صغير. يعتمد القرار الذي نتخذه في هذه الحالة في اختبار فرض العدم على توزيع t وذلك تحت فرض $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (هناك تجانس) وأن كل مجتمع له توزيعاً طبيعياً. أولا نختار عينة عشوائية من الحجم n_1 من المجتمع الأول ونحسب منها $s_1^2, \bar{x}_1$ ونختار عينة عشوائية أخرى من الحجم n_2 من المجتمع الثاني (مستقلة عن العينة الأولي) ونحسب منها $\bar{x}_2$, s_1^2 .

التباين التجميعي pooled variance نحصل عليه من الصيغة التالية :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

من نظرية (٧-١٠) وتحت فرض أن H_0 صحيحاً فإننا نعلم أن :

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

قيمة لمتغير عشوائي T يتبع توزيع t بدرجات حرية $\nu = n_1 + n_2 - 2$ ، في حالة اختبار ذي جانبين وعند مستوي معنوية α فإن منطقة الرفض سوف تكون $T > t_{\frac{\alpha}{2}}$ أو $T < -t_{\frac{\alpha}{2}}$. للبديل من جانب واحد $\mu_1 < \mu_2$ فإن منطقة الرفض سوف تكون $T < -t_\alpha$.

وأخيرا للبديل $\mu_1 > \mu_2$ فإن منطقة الرفض سوف تكون $T > t_\alpha$.

مثال (٩-١٠) اختيرت مجموعتان من الطلبة وأعطيت المجموعة الأولي الوجبة A يوميا أعطيت المجموعة الثانية الوجبة B يوميا . وقد استمرت التجربة لمدة شهر وكانت الزيادة في وزن مفردات كل مجموعة (بالرطل) هي :

المجموعة الأولي	2.6, 2.7, 3.9, 3.4, 1.0, 1.6, 4.0, 3.6, 2.4, 3.0
المجموعة الثانية	2.9, 1.4, 2.6, 1.9, 1.9, 2.4, 2.9, 3.6, 1.6

فهل تعتقد أن هناك فرقا معنوية بين تأثير الوجبة A والوجبة B على زيادة الوزن ؟ (عند مستوى معنوية $\alpha = 0.1$)، وذلك تحت فرض أن العينتين تم اختبارهما من مجتمعين طبيعيين .

الحل .

$$n_1 = 10, \quad \overline{x}_1 = 2.820, \quad s_1 = 0.976$$
$$n_2 = 9, \quad \overline{x}_2 = 2.356, \quad s_2 = 0.716$$

أولا يجب التأكد من أن $\sigma_1^2 = \sigma_2^2$ أى اختبار فرض العدم :

$$H_0 : \sigma_1^2 = \sigma_2^2$$

ضد الفرض البديل :

$$H_1 : \sigma_1^2 \neq \sigma_2^2.$$
$$\alpha = 0.1 .$$

$$f = \frac{s_1^2}{s_2^2} = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}} = \frac{(0.976)^2}{(0.716)^2} = 1.858.$$

$f_{.05}(9,8) = 3.39$ والمستخرجة من جدول توزيع F في ملحق (٦) عنــــد درجــات حرية $\nu_1 = 9, \nu_2 = 8$ أما $f_{.95}(9,8)$ فتحسب من العلاقة التالية :

$$f_{0.95}(9,8) = \frac{1}{f_{0.05}(8,9)} = \frac{1}{3.23} = 0.3096.$$

منطقة الرفض $F > 3.39$ أو $F < 0.3096$

وبما أن f تقع في منطقة القبول فإننا نقبل فرض العدم أن $\sigma_1^2 = \sigma_2^2$.

الآن نختبــر :

$$H_0 : \mu_1 = \mu_2,$$
$$H_1 : \mu_1 \neq \mu_2$$
$$\alpha = 0.1.$$

$$s_p = \sqrt{\frac{s_1^2(n_1-1)+s_2^2(n_2-1)}{n_1+n_2-2}}$$

$$= \sqrt{\frac{(.976)^2(9)+(0.716)^2(8)}{10+9-2}}$$
$$= \sqrt{0.7455548} = 0.86346.$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$
$$= \frac{2.820 - 2.356}{0.86346\sqrt{\frac{1}{10}+\frac{1}{9}}} = 1.16955.$$

$t_{0.05} = 1.74$ والمستخرجة من جدول توزيع t في ملحق (٤) عند درجات حرية $\nu = 17$.

منطقة الرفض $T > 1.74$ أو $T < -174$. وبما أن t تقع في منطقة القبول فإننا نقبــل H_0

وهذا يدل على عدم وجود فرق معنوي بين تأثير الوجبة A والوجبة B على زيادة الوزن .

الحالة الثالثة : عند اختبار فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديــل $H_1 : \mu_1 \neq \mu_2$ تحت الشروط التالية :

(أ) كل مجتمع (من المجتمعين تحت الدراسة) يتبع توزيعاً طبيعياً .

(ب) تباين المجتمعين ، $\sigma_1^2 = \sigma_2^2$ ، مختلفين كثيرا.

(ج) العينتان صغيرتان وإحجامهما مختلفان .

القرار الذي نتخذه يعتمد على الإحصاء T' والذي تقريباً يتبع توزيع t بدرجات حرية تحسب من الصيغة التالية :

$$v = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\left[\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}\right]}$$

لإجراء الاختبار نختار عينة عشوائية حجمها n_1 من المجتمع الأول كما نختار عينة عشوائية أخرى حجمها n_2 من المجتمع الثاني (العينة الثانية مستقلة عن العينة الأولي). نحسب $\bar{x}_1, \bar{x}_2$, s_1^2, s_2^2 وعلى ذلك يكون :

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

قيمة للمتغير T' عندما يكون H_0 صحيحاً. لاختبار ذي جانبين ، بمستوى معنوية α ، منطقة الرفض تقريباً تعطى حيث $t_{\frac{\alpha}{2}}$ أو $-t_{\frac{\alpha}{2}}$ هما القيمتين الحرجتين لتوزيع t بدرجات حرية v و منطقة الرفض سوف تكون $T' > t_{\frac{\alpha}{2}}$ أو $T' < -t_{\frac{\alpha}{2}}$. لبديل من جانب واحد $\mu_1 < \mu_2$ ، فإن منطقة الرفض سوف تكون $T' < -t_\alpha$ وللبديل $\mu_1 > \mu_2$ فإن منطقة الرفض سوف تكون $T' > t_\alpha$.

مثال (٩-١١) أوضحت الدراسة أن زيادة النترات nitrate في الاستهلاك الآدمي له تأثيرات ضارة منها قلة إنتاج الثيروكسين وقلة إدرار اللبن عند البقر. البيانات التالية نتيجة تجربة لقياس النسبة المئوية للزيادة في وزن فئران تجارب صغيرة العمر تناولت وجبة قياسية وفئران تناولت 2000 ppm نترات من مياه الشرب.

النترات	12.7	19.3	20.5	10.5	14.0	10.8	16.6	14.0	17.2
المراقبة (القياسية)	18.2	32.9	10.0	14.3	16.2	27.6	15.7		

تحقق من صحة الفرض القائل : لا يوجد فرق معنوي بين مجموعة النترات ومجموعة المراقبة وذلك عند مستوى معنوية $\alpha = 0.1$. (تحت فرض أن العينتين تم اختيارهما من مجتمعين طبيعيين).

الحل . يجب علينا أولا التحقق من $\sigma_1^2 = \sigma_2^2$.

$s_2 = 8.053$, $s_1 = 3.558$ وعلى ذلك فإن قيمة f هي :

$$f = \frac{s_2^2}{s_1^2} = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}} = \frac{(8.053)^2}{(3.558)^2} = 5.1228.$$

$f_{.05}(6, 8) = 3.58$ والمستخرجة من جدول توزيع F في ملحق (٦) عند درجات حرية $v_1 = 6, v_2 = 8$. أما $f_{.95}(6, 8)$ فيمكن الحصول عليها من العلاقة التالية :

$$f_{.95}(6,8) = \frac{1}{f_{.05}(8,6)} = \frac{1}{4.15} = 0.241.$$

منطقة الرفض $F > 3.58$ أو $F < 0.241$

وحيث أن f تقع في منطقة الرفض فإننا نرفض H_0 ونستنتج أن $\sigma_1^2 \neq \sigma_2^2$

الآن نختبر :

$$H_0 : \mu_1 = \mu_2,$$
$$H_1 : \mu_1 \neq \mu_2.$$
$$\alpha = 0.1.$$
$$\bar{x}_2 = 19.271 \quad , \quad \bar{x}_1 = 15.067$$

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

$$= \frac{15.067 - 19.271}{\sqrt{\frac{3.558^2}{9} + \frac{8.053^2}{7}}} = -1.2869.$$

وعلينا أن نقارن قيمة t' المحسوبة بقيمة t الجدولية عند درجات حرية :

$$\nu = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\left[\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}\right]}$$

$$= \frac{\left[\frac{3.558^2}{9} + \frac{8.053^2}{7}\right]^2}{\left[\frac{\left(\frac{3.558^2}{9}\right)^2}{8} + \frac{\left(\frac{8.053^2}{7}\right)^2}{6}\right]}$$

$$= \frac{113.87}{14.55} = 7.8 \approx 8.$$

$t_{0.05} = 1.86$ والمستخرجة من جدول توزيع t في ملحق (٤) عند درجات حرية $\nu = 8$ وعلى ذلك فإن منطقة الرفض $T' > 1.86$ أو $T' < -1.86$. بما أن t تقع في منطقة القبول نقبل H_0 وهذا يعنى عدم وجود فرق معنوي بين مجموعة النترات ومجموعة المراقبة عند مستوى معنوية $\alpha = 0.1$.

(٩-٨) اختبارات t للأزواج The Paired t Tests

في البند (٩ -٧) كان اهتمامنا بالعينات المستقلة . الآن سوف يكون اهتمامنا بالعينات المزدوجة paired samples ، حيث $d_1, d_2, \dots, d_n$ تمثل الفروق لأزواج المشاهدات المرتبطة التي عددها n . مثل هذه المشاهدات تحدث عندما نأخذ المشاهدات (القراءات) على المفردات قبل وبعد معالجة. بالنظر إلى الفروق لكل أزواج المشاهدات فإننا نأمل في الوصول إلى استنتاج يخص تأثير المعالجة. متوسط فروق المجتمع ، μ_D ، سوف يساوى الفرق بين متوسطي المجتمعين ، $\mu_1 - \mu_2$ ، وعلى ذلك فإن مشكلة اختبار فرض العدم H_0 أن $\mu_1 - \mu_2 = 0$ تكافئ اختبار $\mu_D = 0$. بفرض أن كل مجتمع من المجتمعين يتبع توزيعاً طبيعياً. أولاً نختار n

من أزواج المشاهدات عشوائيا ونحسب الفروق ونقدر $\bar{d}$ و s_d. وعلى ذلك تبعا لنظرية (٧-١١) ، فإننا نعلم أنه عندما يكون H_0 صحيحا فإن :

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

هي قيمة للمتغير العشوائي T الذي يتبع توزيع t بدرجات حرية $v = n - 1$. قرارنا في هذه الحالة يعتمد على توزيع t ومناطق الرفض المقابلة للفروض البديلة المختلفة وسوف نتبع الخطوات التي أتبعت من قبل .

مثال (٩-١٢) إذا كانت أوزان 10 أشخاص قبل التوقف عن التدخين وبعد 8 أسابيع من التوقف عن التدخين كما يلي :

الأشخاص	1	2	3	4	5	6	7	8	9	10
قبل	148	176	152	115	166	147	150	175	151	146
بعد	155	170	169	117	170	150	154	180	160	151

فهل تدل هذه البيانات على أن الامتناع عن التدخين يؤدى إلى زيادة وزن الأشخاص الذين يمتنعون عن التدخين ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل .

$$H_0 : \mu_D = 0,$$
$$H_1 : \mu_D \neq 0.$$
$$\alpha = 0.05.$$

$t_{0.025} = 2.262$ والمستخرجة من جدول توزيع t في ملحق (٤) عند درجات حرية $v = 9$.

منطقة الرفض $T > 2.262$ أو $T < -2.262$

$$\bar{d} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{-50}{10} = -5 \quad ,$$

$$s_d = \sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{n} d_i^2 - \frac{(\sum_{i=1}^{n} d_i)^2}{n}\right]}$$

$$= \sqrt{\frac{1}{9}\left[550 - \frac{(-50)^2}{10}\right]} = 5.7735,$$

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

$$= \frac{-5}{5.7735/\sqrt{10}} = -2.7386.$$

وبما أن t تقع في منطقة الرفض نرفض H_0

(٩-٩) اختبارات تخص نسبة مجتمع

Tests Concerning a Population Proportion

سوف نهتم في هذا البند بمشكلة اختبارات الفروض التي فيها نسبة صفة ما تساوى قيمة معنية . أى أننا نهتم باختبار فرض العدم $H_0 : p = p_0$ ضد الفرض البديل $p < p_0$ أو $p > p_0$ أو $p \neq p_0$. سوف تقتصر دراستنا في هذا البند على حالة العينات الكبيرة ، وعلى ذلك فإن الإحصاء المناسب الذي يعتمد عليه قرارنا هو $\hat{P}$ الذي تقريباً يتبع توزيعاً طبيعياً. أى أن قرارنا سوف يعتمد على :

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \, q_0}{n}}}$$

والذي يمثل قيمة للمتغير Z الذي يتبع التوزيع الطبيعي القياسي. وعلى ذلك للاختبار من جانبين فإن منطقة الرفض ، بمستوى معنوية α، سوف تكون $Z > z_{\frac{\alpha}{2}}$ أو $Z < -z_{\frac{\alpha}{2}}$. للبديل من جانب واحد $p < p_0$ فإن منطقة الرفض سوف تكون $Z < -z_{\alpha}$. وأخيرا للبديل $p > p_0$ فإن منطقة الرفض سوف تكون $Z > z_{\alpha}$.

مثال (٩ – ١٣) يدعى منتج أن 90% من قطع الغيار التي يمد بها مصنعاً مطابقاً للمواصفات . فإذا تم اختيار عينة عشوائية من 200 قطعة ووجد أن 40 قطعة تالفة. أختبر أدعاء المنتج عند مستوى معنوية $\alpha = 0.05$.

$$H_0 : \; p = 0.9,$$
$$H_1 : \; p \neq 0.9.$$
$$\alpha = 0.05.$$

$z_{.025} = 1.96$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣). منطقة الرفض $Z > 1.96$ أو $Z < -1.96$

$$x = 160 \quad , \quad n = 200.$$

$$\hat{p} = \frac{x}{n} = \frac{160}{200} = 0.8 \ , \ q_0 = 1 - q_0 = 1 - 0.9 = 0.1.$$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \ q_0}{n}}} = \frac{0.8 - 0.9}{\sqrt{\frac{(0.9)(0.1)}{200}}} = -4.714.$$

وبما أن z تقع في منطقة الرفض نرفض H_0 .

(٩-١٠) اختبارات تخص الفرق بين نسبتي مجتمعين

Tests Concerning a Difference Between Two Population Proportions

بفرض أن p_1 هي نسبة توفر صفة ما في إحدى المجتمعات وكانت p_2 هي نسبة توفر الصفة نفسها في مجتمع آخر وإذا كان اهتمامنا باختبار فرض العدم $H_0 : p_1 = p_2$ فإن الإحصاء المناسب والذي يعتمد عليه قرارنا سوف يكون المتغير العشوائي $\hat{P}_1 - \hat{P}_2$. نختار عينة عشوائية كبيرة من الحجم n_1 من المجتمع الأول ونحسب نسبة توفر الصفة محل الدراسة فيها ولتكن $\hat{p}_1 = \frac{x_1}{n_1}$ حيث أن x_1 هي عدد الذين يمثلون الصفة في المجتمع الأول. نختار عينة عشوائية أخرى كبيرة من الحجم n_2 من المجتمع الثاني ونحسب نسبة توفر الصفة المطلوبة منها ولتكن $\hat{p}_2 = \frac{x_2}{n_2}$ حيث x_2 هي عدد الذين يمتلكون الصفة في المجتمع الثاني ، ويجب أن تكون العينتين مستقلين تحت فرض العدم ومن نظرية (٧-٨) فإن :

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 \ q_1}{n_1} + \frac{p_2 \ q_2}{n_2}}}$$

$$= \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq[\frac{1}{n_1} + \frac{1}{n_2}]}}.$$

هو قيمة لمتغير عشوائي Z يتبع التوزيع الطبيعي القياسي عندما H_0 يكون صحيحا و n_2 , n_1 كبيرتان . وبما أن p مجهولة في صيغة z فإننا نحسبها من الصيغة التالية :

$$\tilde{p} = \frac{x_1 + x_2}{n_1 + n_2}.$$

وعلى ذلك تصبح z كالتالي :

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\tilde{p}\ \tilde{q}[\frac{1}{n_1} + \frac{1}{n_2}]}}.$$

حيث أن $\tilde{q} = 1 - \tilde{p}$

منطقة الرفض للفروض البديلة المختلفة يمكن الحصول عليها ، كما سبق أن ذكرنا ، باستخدام القيم الحرجة لمنحنى التوزيع الطبيعي القياسي .

مثال (٩ – ١٤) اختبرت عينة عشوائية من 300 مدخنا في مدينة ما ووجد أن من بينهم 60 يفضلون تدخين النوع A من السجائر ثم اختبرت عينة عشوائية من 200 مدخنا في مدينة أخرى ووجد أن من بينهم 30 يفضلون تدخين النوع A من السجائر . اختبر فرض العدم $H_0 : p_1 = p_2$ ضد الفرض البديل $H_1 : p_1 \neq p_2$ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل •

$$H_0 : p_1 = p_2,$$
$$H_1 : p_1 \neq p_2.$$
$$\alpha = 0.05.$$

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{60}{300} = 0.2 \quad , \quad \hat{p}_2 = \frac{x_2}{n_2} = \frac{30}{200} = 0.15,$$

$$\tilde{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{60 + 30}{300 + 200} = \frac{90}{500} = 0.18, \tilde{q} = 1 - \tilde{p} = 1 - 0.18 = 0.82.$$

$z_{.025} = 1.96$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣).

منطقة الرفض $Z > 1.96$ أو $Z < -1.96$

$$z = \frac{(0.2 - 0.15) - 0}{\sqrt{(0.18)(0.82)[(\frac{1}{300}) + (\frac{1}{200})]}} = 1.4257.$$

بما أن z تقع في منطقة القبول فإننا نقبل H_0 .

<u>تمـــارين :</u>

-١- بين أي الجمل التالية صواب وأيهم خطأ .

أ- $H_0 : S \leq 0.2$ ب- $H_0 : \bar{x} = 45$

ج- $H_0 : \mu = 100\ ,\ H_1 : \mu > 100$

د- $H_0 : \overline{X}_1 - \overline{X}_2 = 5$

هـ $H_0 : p = 0.25 \; , \; H_1 : p \neq 0.25$

و- $H_0 : S_1^2 = S_2^2 \; , \; H_1 : S_1^2 \neq S_2^2$

-٢- إذا كانت μ تمثل المتوسط الحقيقي لمستوى الإشعاع (مقاسه Picocuries per liter). القيمة 5 pci/L تمثل القيمة الحرجة (الحد الفاصل بين الأمان وعدم الأمان للماء) هل تقترح فرض العدم $H_0 : \mu = 5$ ضد الفرض البديل $H_1 : \mu > 5$ أو اختبار فرض العدم $H_0 : \mu = 5$ ضد الفرض البديل $H_1 : \mu < 5$ ولماذا ؟

-٣- إذا كان الإحصاء Z يتبع توزيع طبيعي قياسي عندما تكون H_0 صحيح أوجد مستوى المعنوية لكل من الحالات التالية :

أ- $H_1 : \mu > \mu_0$ ومنطقة الرفض $Z > 1.88$.

ب- $H_1 : \mu < \mu_0$ ومنطقة الرفض $Z < -2.75$.

ج- $H_1 : \mu \neq \mu_0$ ومنطقة الرفض $Z > 2.88$ أو $Z < -2.88$.

-٤- مجتمع طبيعي له متوسط μ وانحراف معياري $\sigma = 0.5$ ، لاختبار فرض العدم $H_0 : \mu = 15$ اختيرت عينة عشوائية من الحجم $n = 25$ وكانت منطقة الرفض $\overline{X} < 14.9$ أوجد :

أ- احتمال الوقوع في خطأ من النوع الأول ؟

ب- احتمال الوقوع في خطأ من النوع الثاني إذا كان الفرض البديل $\mu = 14.8$ و $\mu = 14.9$.

-٥- تنص المواصفات القياسية على أن قوة الضغط Compression Strength لخليط من الأسمنت البورتلاندى ومركبات أخرى لابد أن تزيد عن 1300 k N/m . بفرض أن قوة الضغط لهذا الخليط يتبع توزيعاً طبيعياً بمتوسط μ وانحراف معياري $\sigma = 60$.

(ا) ما هو فرض العدم والفرض البديل المناسب لاختبار μ في هذه الحالة ؟

(ب) إذا كانت $\overline{X}$ تمثل الإحصاء المناسب واختيرت عينة عشوائية من الحجم $n = 25$ وإذا كانت منطقة الرفض $\overline{X} \geq 1331.26$. ما هو احتمال الوقوع في خطأ من النوع الأول ؟

-٦- إذا كانت النسبة المئوية المرغوبة لمركب S_iO_2 في نوع معين من الأسمنت (في المتوسط) هي 5.5 ، لاختبار فرض العدم $H_0 : \mu = 5.5$ اختيرت 16 عينة مستقلة للتحليل . بفرض أن النسبة المئوية للمركب S_iO_2 في العينة تتبع توزيعاً طبيعياً بانحراف معياري $\sigma = 0.3$

وإذا كانت $\bar{x} = 5.25$. اختبر فرض العدم $H_0 : \mu = 5.5$ ضد الفرض البديل $H_1 : \mu \neq 5.5$ عند مستوى معنوي $\alpha = 0.01$.

-٧- يتطلب العمليات الجراحية التي تجري للخيول في ظروف الحقل مخدر يؤدى إلى تخدير الحيوان لزمن محدد. اختيرت عينة عشوائية من 73 حيوانا كانوا تحت المخدر وتم حساب أزمنة التخدير لهم وكان $\bar{x} = 18.86$ دقيقة بانحراف معياري $s = 8.6$ دقيقة هل هذه البيانات تدل على أن متوسط زمن الغياب عن الوعي تحت ظروف الحقل أقل من 20 دقيقة ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨- يوصي معهد التغذية على أن كمية الزنك التي يحتاج إليها الفرد الذكر في العمر أكبر من 50 سنة هو 15 mg\ day أجريت تجربة وتم الحصول على النتائج التالية :

$\bar{x} = 11.3, \quad n = 115, \quad s = 6.45$

فهل تدل هذه البيانات على أن متوسط الاحتياج اليومي من الزنك في مجتمع الذكور من العمر أكبر من 50 تقل عن المسموح به ؟

-٩- تقوم شركة لصناعة الإطارات بإنتاج نوع من إطارات السيارات التي يتحمل ضغط $30\ lb / in^2$. فإذا كانت μ هو المتوسط الحقيقي للضغط . أوجد مستوى المعنوية α لقيم z التالية :

أ- 2.10 ب - 1.75 ج- 1.41 د-5.3

-١٠- إذا كانت μ تمثل متوسط سير الدم لكل النساء الحوامل . لاختبار فرض العدم $H_0 : \mu = 5.63$ ضد الفرض البديل $H_1 : \mu \neq 5.63$ وذلك بالاعتماد على عينة عشوائية من النساء الحوامل من الحجم $n = 176$. فهل تعتقد أنه يمكن رفض H_0 عند مستوى معنوية $\alpha = 0.01$ إذا كانت $\bar{x} = 6.1$, $s = 0.7$ ؟

-١١- تقوم شركة لتصنيع الأدوية بإنتاج نوع معين من الأسبرين في عبوات سعة العبوة 100 حبة. تهتم الشركة بوزن الحبة أكثر من اهتمامها بالعدد ، فإذا كان من الضروري أن يكون متوسط وزن الحبة 5 grains . اختيرت عينة عشوائية من الحجم $n = 100$ حبة من إنتاج كبير وكان متوسط وزن الحبة $\bar{x} = 4.87$ بانحراف معياري $s = 0.35$. هل المعلومات التي تم الحصول عليها كافية لاستنتاج أن الشركة تنتج طبقاً للمواصفات القياسية ؟ اختبر الفرض المناسب عند مستوى معنوية 0.01 .

-١٢- أوضحت الخبرة الماضية أن درجات الطلبة في مادة الإحصاء تتبع توزيعاً طبيعياً بمتوسط 75 وتباين 16 . يرغب الأعضاء في قسم الرياضيات في معرفة هل متوسط درجات الطلبة في

السنة الحالية لها نفس مستوى السنوات الماضية ؟ لذلك قرروا اختبار فرض العدم $H_0 : \mu = 75$ ضد الفرض البديل $\mu \neq 75$ عند مستوى معنوية $\alpha = 0.05$. اختيرت عينة عشوائية من الحجم $n = 15$ وكان متوسط درجات الطلبة في العينة $\bar{x} = 80$. ما الاستنتاج الذي يمكن وضعه ؟

-١٣- تسلم أحد التجار كمية كبيرة من بطاريات السيارات المنتجة بواسطة مصنع جديد. يعتقد مدير المصنع أن متوسط عمر البطارية المنتجة 30 شهرا . لاختبار ذلك اختيرت عينة عشوائية من البطاريات تسلمها التاجر وتمت تجربتها فكانت أعمارها بالشهر كالآتي:

29.9, 30.0, 30.2, 35.4, 37.6, 37.6, 92.6, 29.8, 34.7, 39.8, 35.4.

فإذا كانت أعمار البطاريات المنتجة في المصنع تتبع توزيعاً طبيعياً فهل تدل بيانات العينة على أن متوسط أعمار البطاريات أقل من 30 شهر إذا علم أن الانحراف المعياري للبطاريات المنتجة في هذا المصنع $\sigma = 3$ شهرا .

-١٤- إذا كان متوسط الذكاء لعينة عشوائية ، من الحجم $n = 20$ ، هو $\bar{x} = 1053$. وبفرض أن درجات الذكاء في المجتمع الذي اختيرت فيه العينة يتبع توزيعاً طبيعياً بانحراف معياري 3.7 .تحقق من صحة الفرض القائل أن متوسط ذكاء المجتمع الذي اختيرت منه العينة لا يختلف عن 106.2 عند مستوى معنوية $\alpha = 0.05$.

-١٥- اختيرت عينة عشوائية من 25 عاملا بإحدى الشركات وكان متوسط إنتاج العامل في العينة 28 وحدة في اليوم، علما بأن إنتاجية العامل في هذه الشركة تتبع توزيعاً طبيعياً بتباين $\sigma^2 = 5$ اختبر الفرض القائل أن متوسط إنتاجية العامل اليومية في هذه الشركة 25 وحدة وذلك عند مستوى معنوية $\alpha = 0.01$.

-١٦- إذا كان من المعروف أن كمية المحصول بالمنتج من إحدى المحاصيل يتبع توزيعاً طبيعياً بانحراف معياري 30 إردب للفدان ، اختيرت عينة عشوائية مساحتها 15 فدانا وتم حساب متوسط المحصول في العينة فكانت 490 إردب للفدان أختبر الفرض القائل أن الكمية المتوقعة للمحصول تساوى 500 إردب للفدان وذلك عند مستوى معنوية $\alpha = 0.05$.

-١٧- يدعى مسئول في مصنع لإنتاج المصابيح الكهربائية أن متوسط عمر المصابيح هو 1000 ساعة، مع العلم أن أعمار المصابيح من إنتاج هذا المصنع يتبع توزيعاً طبيعياً بانحراف معياري 30 ساعة . اختيرت عينة عشوائية من الحجم $n = 25$ من إنتاج هذا المصنع فكان $\bar{x} = 900$ فهل تدل هذه البيانات أن متوسط عمر المصابيح 1000 ساعة وذلك عند مستوى معنوية $\alpha = 0.01$ •

-١٨- تستخدم آلة لملئ أكياس بسلعة ما بطريقة أتوماتيكية بمتوسط 55 جرام للكيس. اختيرت عينة عشوائية من 35 كيس من هذه الآلة وتم حساب متوسط وزن الكيس فكان $\bar{x} = 57$ والانحراف المعياري $s = 3$ جرام أختبر فرض العدم $H_0 : \mu = 55$ ضد الفرض البديل $H_1 : \mu \neq 55$ وذلك عند مستوى معنوية $\alpha = 0.05$.

-١٩- للمعلومات التالية إذا علم أن العينة تم اختيارها من مجتمع له توزيعاً طبيعياً ، أوجد منطقة الرفض وأحسب قيمة z وقرر هل تقبل فرض العدم أم لا ؟

$$H_0 : \mu = 30, \quad H_1, \mu < 30, \quad \sigma = 6, \quad , n = 20, \quad \alpha = 0.05, \quad \bar{x} = 25.$$

-٢٠- يعتقد مسئول في دار للنشر أنه في المتوسط تطبع الآلة 45 نسخة في الدقيقة من مجلد ما علما بأن المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً بانحراف معياري $\sigma = 2.1$ دقيقة . في محاولة لزيادة عدد النسخ التي تطبع تم إدخال تعديلات على آلة الطبع ثم أخذت عينة عشوائية من الحجم $n = 25$ فكان متوسط عدد النسخ في الدقيقة $\bar{x} = 55$.اختبر فرض العدم $H_0 : \mu = 45$ ضد الفرض البديل $\mu > 45$ عند مستوى معنوية $\alpha = 0.05$ •

-٢١- إذا كان متوسط القامة في مجتمع ما هو 172 سم . اختيرت عينة عشوائية من هذا المجتمع فكان متوسط الطول فيها 173 سم بانحراف معياري 3.5 . فإذا كان حجم العينة 30 شخصا • اختبر فرض العدم $\mu = 172$ ضد الفرض البديل $\mu \neq 172$ عند مستوى معنوية $\alpha = 0.01$.

-٢٢- اختيرت.عينة عشوائية من 35 أسرة وحسب متوسط الدخل الشهري للأسرة فوجد أنه $500 بانحراف معياري 40 $. اختبر فرض العدم أن متوسط دخل الأسرة في هذا المجتمع يساوى 450 $ ضد الفرض البديل $\mu \neq 450$ وذلك عند مستوى معنوية $\alpha = 0.01$.

-٢٣- في دراسة لمعرفة تأثير نوع معين من معجون الأسنان على مقاومة تسوس الأسنان اختيرت عينة عشوائية من 25 طالبا وأشرف عليهم أحد المدرسين حتى يستعملوا هذا المعجون يوميا ثم قيست المقاومة بعد ثلاثة أشهر من هذه المحاولة بمقياس معين لعدد الميكروبات من نوع معروف موجود في اللعاب فكانت $\bar{x} = 5$.فإذا كان معروف للباحث أن هذا المقياس يتبع التوزيع الطبيعي بانحراف معياري $\sigma = 0.49$. فهل هناك ما يدعو إلى الاعتقاد بأن لهذا المعجون فائدة على زيادة المقاومة ضد التسوس وذلك عند مستوى معنوية $\alpha = 0.01$ ؟علما بأن $\mu_0 = 3$.

-٢٤- قررت شركة ما لا تزيد مدة المكالمة التليفونية التي يطلبها الموظف عن 25 دقيقة . اختيرت عينة عشوائية من 35 مكالمة فأعطيت متوسط 26.25 دقيقة بانحراف معياري 2.1

دقيقة ما الاستنتاج الذي يمكن اتخاذه بناء على هذه العينة وذلك عند مستوى معنوية $\alpha = 0.05$.

-٢٥- اختيرت عينة من 35 عاملا في مصنع وكان متوسط العمر 38 سنة بانحراف معياري 8 سنة. أختبر فرض العدم $\mu = 40$ ضد الفرض البديل $\mu \neq 40$ عند مستوى معنوية $\alpha = 0.05$.

-٢٦- في تقرير خاص وجد أن كمية مركب **tar** لكل سيجارة من نوع ما هو 13 ملليجرام . اختيرت عينة عشوائية من 100 سيجارة فأعطيت متوسط 12.85 من **tar** بانحراف معياري 0.75 اختبر فرض العدم $H_0 : \mu = 13$ ضد الفرض البديل $H_1 : \mu \neq 13$ عند مستوى معنوية $\alpha = 0.01$.

-٢٧- قام مسئول من مصنع للعب الأطفال يعمل تعديل في سيارات الأطفال وكان من نتيجة هذا التعديل زيادة زمن عمل البطارية المشغلة للسيارة اختبرت عينة عشوائية من 100 سيارة وتم حساب متوسط عمر البطارية فكان 31 ساعة بانحراف معياري 0.75 اختبر فرض العدم $H_0 : \mu = 25$ وذلك عند مستوى معنوية $\alpha = 0.01$.

-٢٨- إذا كانت نقطة الذوبان المحسوبة من 16 عينة من نوع ما من زيت الطعام المهدرج هى $\bar{x} = 94.32$ بفرض أن توزيع نقطة الذوبان طبيعي بانحراف معياري $\sigma = 1.2$. المطلوب اختبار فرض العدم $H_0 : \mu = 95$ ضد الفرض البديل $H_1 : \mu \neq 95$ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٢٩- يعتقد مسئول في مصنع للتليفزيونات أن التيار اللازم للحصول على صورة واضحة لشاشة التليفزيون على الأكثر (μA) 250 **microamperes** اختيرت عينة عشوائية من 20 وحدة (تليفزيون) وكان متوسط التيار من العينة $\bar{x} = 257.3$ فإذا كان μ تمثل متوسط التيار الحقيقي الضروري للحصول على الصورة الواضحة للتليفزيون من هذا النوع وتحت فرض أن μ هى متوسط توزيع طبيعي بانحراف معياري $\sigma = 15$ ، اختبر فرض العدم $H_0 : \mu = 250$ ضد الفرض المناسب وذلك عند مستوى معنوية $\alpha = 0.01$.

-٣٠- إذا كان T متغيرا عشوائياً يتبع توزيع t عندما يكون H_0 صحيح أوجد مستوى المعنوية α لكل من الحالات الآتية : (درجات الحرية = ν) .

أ- $\nu = 15$, $H_1 : \mu > \mu_0$

ومنطقة الرفض $T \geq 3.733$

ب- $n = 24$, $H_1 : \mu < \mu_0$

ومنطقة الرفض $T \leq -2.069$

ج- $H_1 : \mu \neq \mu_0 \;,\; n = 31$

ومنطقة الرفض $T > 1.697$ $T \leq -1.697$

-٣١- يقوم باحث بجمع البيانات لاختبار فرض العدم $H_0 : \mu = 17$ ضد الفروض البديل $H_1 : \mu > 17$ أوجد مستوى المعنوية α المرتبط بقيم t ودرجات الحرية التالية :

أ- $\nu = 14 \;,\; t = 1.761$

ب- $\nu = 25 \;,\; t = 3.450$

ج- $\nu = 13 \;,\; t = 1.771$

د- $\nu = 8 \;,\; t = 1.860$

هـ $\nu = 40 \;,\; t = 1.684$

-٣٢- أوجد مستوى المعنوية α لاختبار من جانبين في الحالات التالية :

أ- $\nu = 6 \;,\; t = 2.447$

ب- $\nu = 14 \;,\; t = 2.145$

ج- $\nu = 24 \;,\; t = 2.064$

د- $\nu = 17 \;,\; t = 2.110$

-٣٣- اختيرت عينة عشوائية مكونة من درجات 10 طلاب في مادة الإحصاء وكان $\bar{x} = 7 \;,\; s = 3$. أختبر الفرض القائل أن متوسط درجة الطالب في المجتمع الطبيعي المسحوب منه العينة يساوى 7.5 وذلك عند مستوى معنوية $\alpha = 0.05$.

-٣٤- اختيرت عينة عشوائية من 10 فكانت أطوالهم بالبوصة هى

81, 70, 68, 68, 64, 72, 71, 80, 61, 70

اختبر الفرض القائل أن $\mu = 75$ ضد الفرض البديل $\mu \neq 75$ عند مستوى معنوية $\alpha = 0.01$. مع العلم بأن المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً .

-٣٥- يعتقد مدير شركة لصناعة الصابون أن أوزان صناديق الصابون يتبع التوزيع الطبيعي بمتوسط 500 جرام . اختيرت عينة عشوائية من 15 صندوقا ووجد أن $\bar{x} = 530$ و $s = 5.5$ جراما . اختبر فرض العدم أن $\mu = 500$ ضد الفرض البديل $\mu \neq 500$ عند مستوى معنوية $\alpha = 0.05$.

-٣٦- اختيرت عينة عشوائية من 20 عبوة من مشروب بارد استخدمت آلة لتعبئته. فإذا كان متوسط العبوة $\bar{x} = 7.5$ أوقية بانحراف معياري 0.47 أوقية. اختبر فرض العدم أن $\mu = 7.8$

ضد الفرض البديل $\mu < 7.8$ عند مستوى معنوية $\alpha = 0.01$ تحت فرض أن المجتمـــع الـــذي اختيرت منه العينة يتبع توزيعاً طبيعياً .

-٣٧- لاختبار فرض العدم أن متوسط وزن الصندوق من القمح المعبأ في شركة ما هـــو 10 كيلو . اختيرت عينة عشوائية من 10 صناديق فكانت النتائج كالتالي :

10.1, 9.5, 10.1, 11.2, 9.9, 8.7, 6.7, 8.1, 9.9, 6.1

استخدم مستوى معنوية $\alpha = 0.01$ لاختبار فرض العدم وذلك تحت فرض أن توزيـــع الأوزان لصناديق القمح طبيعي .

-٣٨- يرغب باحث في أن تكون حموضة التربة التي سوف يستخدمها في أبحاثه ذات حموضـــة PH= 8.75 . لهذا الفرض قام الباحث بمعالجة التربة بمحلول من 25% من الماء و 75% مـــن effluent وبعد ذلك تم اختيار 5 عينات مشتقة من التربة المعالجة فكـــان المتوســـط الحســـابي والانحراف المعياري هما $s = 0.05$, $\bar{x} = 8$ فهل تدل هذه البيانات علـــى أن حموضـــة التربـــة تساوى 8.75 عند مستوى معنوية $\alpha = 0.05$ مع العلم أن العينة تم اختيارها من مجتمع طبيعي .

-٣٩- لمثال (٣٨) اختبر فرض العدم $H_0 : \sigma = .55$ ضد الفرض البديل $H_1 : \sigma \neq .55$ عند مستوى معنوية $\alpha = 0.05$.

-٤٠- للمثال (٢٧) اختبر فرض العدم $H_0 : \sigma = 0.81$ ضد الفرض البديل $H_1 : \sigma < 0.81$ عند مستوى معنوية $\alpha = 0.01$.

-٤١- قام المسئولون في شركة لإنتاج ملابس الأطفال بإنتاج نوعية من الملابس المقاومة للحريق . للمقارنة بين النوعين تم الحصول على البيانات التالية :

$$n_1 = 15, \quad n_2 = 15, \alpha = 0.05, \quad s_1 = 1.5, \quad s_2 = .2$$

اختبر فرض العدم $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عند مســـتوى معنوية $\alpha = 0.01$ مع العلم أن أزمنة الحريق للملابس في المصنع تتبع توزيعاً طبيعياً .

-٤٢- أجريت دراسة في إحدى مراكز العلاج الطبيعي على عينة عشوائية من 10 أشـــخاص ممن يتبعون نظام إنقاص الوزن وقد تم تسجيل مقدار النقص في الوزن لكل شخص في العينة خلال فترة إتباع النظام المتبع لإنقاص الوزن وتم الحصول على البيانات التالية :

14, 5, 5, 11, 12, 17, 7, 3, 4, 9

اختبر فرض الدم $H_0 : \sigma^2 = 10$ ضد الفرض البديل $H_1 : \sigma^2 > 10$ عند مســـتوى معنوية $\alpha = 0.01$ وذلك تحت فرض أن المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً .

-٤٣- اختيرت عينة عشوائية من الحجم $n = 10$ من مجتمع طبيعي وكان تباين العينة $s^2 = 24$ اختبر فرض العدم $H_0 : \sigma^2 = 23$ حيث الفرض البديل $H_1 : \sigma^2 \neq 23$ وذلك عند مستوى معنوية $\alpha = 0.01$.

-٤٤- في إحدى مراكز تعليم غير المبصرين يوجد نظامين B , A لتعليم القراءة فإذا كان تباين العينة لمستوى القراءة بالنسبة للدارسين عن طريق النظام A هو $s_1^2 = 1.2$ وذلك من عينة عشوائية حجمها $n = 30$. اختيرت عينة عشوائية أخرى مكونة من 35 فردا ممن يدرسون باستخدام النظام B فكان تباين العينة $s_2^2 = 1.04$. اختبر الفرض القائل بأن التدريس باستخدام النظام A متساوي في التشتت مع النظام B عند مستوى معنوية $\alpha = 0.01$.

-٤٥- في عينة عشوائية حجمها $n = 20$ وجد أن الانحراف المعياري لتركيز الصوديوم في الدم (m E q / L) هو $s_1 = 40.5$ بينما في عينة أخرى من الحجم $n_2 = 20$ وجد أن $s_2 = 32.1$. بفرض أن توزيع كمية الصوديوم طبيعي اختبر فرض العدم $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.01$.

-٤٦- يعتبر التوكسافين من المبيدات التي تلوث البيئة ويعتبر خطرا على النبات والحيوان والإنسان. قام باحث باختيار عينة عشوائية من الفئران (الإناث) من الحجم $n_1 = 30$ وتم إعطائهم خلطة غذائية بها جرعة مخفضة من التوكسافين (4 ppm) وأخذت عينة أخرى من الحجم $n_2 = 23$ (اعتبرت عينة المراقبة لم تأخذ أى مبيد في الخلطة الغذائية). وفي نهاية التجربة تم تسجيل الزيادة في الوزن لكل فأر، فإذا كان الانحراف المعياري لعينة المراقبة $s_1 = 32$ جراما بينما الانحراف المعياري للعينة المعالجة $s_2 = 54$ جراما أختبر فرض العدم $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.01$، مع العلم أن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٤٧- اختيرت 10 عينات من الموقع A و 8 عينات من الموقع B وتم قياس الحموضة لكل عينة PH . والبيانات في الجدول التالي :

الموقع A	8.53	8.52	8.01	8.01	7.88	7.93
	9.98	7.85	7.92	7.80		
الموقع B	7.85	7.73	7.58	7.4	3.35	
	7.3	7.27	7.27			

أ- أختبر فرض العدم $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$.

ب- هل هناك فرق معنوي بين متوسطي الحموضة للموقعين عند مستوى معنوية $\alpha = 0.02$
مع العلم أن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٤٨- زرع صنفان من القمح ، الأول من 10 قطع فكان تباين المحصول هو $s_1^2 = 25$ وزرع الصنف الثاني في 15 قطعة وكان تبـــاين المحصـــول هـــو $s_2^2 = 15$ أختـــبر فـــرض العـــدم $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$ مع العلـم أن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٤٩- إذا كانت μ_1 تمثل العمر الحقيقي لإطارات السيارات من النوع A مقاسة بالأميال (عدد الأميال التي تقطعها السيارة حتى يستهلك الإطار) و μ_2 تمثل العمر الحقيقي لإطـــارات السيارات مـــن النـــوع B . اختـــبر فـــرض العدم $H_0 : \mu_1 = \mu_2$ ضـــد الفـــرض البديـــل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ إذا كانت :

$$n_1 = 40, \quad \bar{x}_1 = 36500 \ , \ s_1 = 220,$$
$$n_2 = 40, \quad \bar{x}_2 = 33400 \ , \ s_2 = 190$$

مع العلم بأن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٥٠- طبق اختبار للعصابيه على مجموعتين ، الأولى من الذكور وحجمها 35 والثانيـــة مـــن الإناث وحجمها 40 فإذا كان متوسط العصابية لدي الذكور 21.3 بـــانحراف معيـــاري 4.6 ومتوسط العصابية لدى الإناث 24.2 بانحراف معياري 3.9 تحقق من صحة الفرض القـــائل أن $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$.

-٥١- في دراسة لتقدير ما إذا كان هناك فرق معنوي بين أجور أعضـــاء هيئـــة التدريـــس في جامعتين A , B اختيرت عينة عشوائية من 100 عضو هيئة تدريس من الجامعة A ووجـــد أن $\bar{x}_1 = 11000$ \$ خلال ٩ أشهر بانحراف معياري $s_1 = 1300$. كما اختيرت عينة عشـــوائية أخرى من 200 عضو هيئة تدريس من الجامعة B ووجد أن $\bar{x}_2 = 11900$ \$ بانحراف معيـــاري $s_2 = 1400$. اختبر الفرض القائل أن متوسط الأجور لأعضاء هيئة التدريس خلال 9 اشـــهر في الجامعة A لا يختلف عن متوسط الأجور في الجامعة B عند مستوى معنوية $\alpha = 0.05$.

-٥٢- طبق اختبار لطلاقة الكلمات على مجموعتين الأول من الانبساطين والثانية من العصابيين حجمهما 45 , 50 شخصا على التوالي ، وحصلت مجموعة الانبساطين على متوسط قدرة 70.2 بانحراف معياري 11.6 وحصلت العصابيين على متوسط قدرة 65 بـــانحراف معيـــاري 10.2 . تحقق من صحة الفرض القائل أن طلاقة الكلمات واحدة في المجموعتين؟

٥٣- يعتقد المسئول في مصنع عن وجود اختلاف في جودة نوعين من قطع الغيار تم استلامهم من موردين B , A وقد تم الحصول على البيانات التالية بناء على عينتين عشوائيين من الموردين :

المورد A $n_1 = 50, \quad \bar{x}_1 = 153, \quad s_1 = 10$

المورد B $n_2 = 100, \quad \bar{x}_2 = 150, \quad s_2 = 5$

اختبر القول القائل بعدم وجود فرق معنوي بين العينتين عند مستوى معنوية $\alpha = 0.05$.

٥٤- في إحدى مراكز رعاية الطفل سجلت الأوزان (بالرطل) لعينتين عشوائيتين من الأطفال حديثي الولادة كل منهما يتكون من خمسة أطفال فكانت البيانات كمايلي :

المجموعة الأولي 8, 6, 6, 5, 7

المجموعة الثانية 8, 7, 8, 5, 6

أ- اختبر الفرض القائل أن $\sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $\sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.1$. مع العلم أن العينتين تم اختيارهما من مجتمعين طبيعيين .

ب- أختبر فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.02$.

٥٥- يرغب مسئول في مصنع لانتاج معجون للأسنان في دراسة تأثير إضافة مادة كيمائية معينة إلى معجون لتحسين مفعوله. اختبرت عينتين مستقلتين كل عينة من 10 أشخاص استخدمت أفراد العينة الأولي المعجون مضاف إليه المادة الكيمائية بينما استخدمت أفراد العينة الثانية المعجون بدون إضافة المادة الكيمائية باستخدام مقياس خاص تم الحصول على البيانات التالية :

العينة الأولى (أضيفت المادة الكيمائية) $\bar{x}_1 = 8, \quad s_1 = 3$

العينة الثانية (بدون إضافة) $\bar{x}_2 = 9, \quad s_2 = 4$

اختير فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.01$. مع العلم أن العينتين تم اختيارهما من مجتمعين طبيعيين .

٥٦- مجموعتان من الطلبة تتكون إحداهما من 10 أفراد والأخرى من 8 أفراد أعطيت امتحانا واحدا ودونت النتائج كما يلي :

المجموعة الأولى 7, 7, 7, 6, 7, 8, 8, 9, 9, 10

المجموعة الثانية *10, 10, 8, 7, 7, 10, 9, 7*

المطلوب اختبار الفرض $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.1$. مع العلم أن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٥٧- في اختبار للقلق لمجموعة من الذكور والإناث تم الحصول على البيانات التالي :

	حجم العينة	متوسط العينة	الانحراف المعياري
ذكـور	20	10.4	4.83
إناث	25	9.26	4.68

اختبر فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ مع العلم بأن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٥٨- البيانات التالية تمثل أوزان الأدوات الخاصة بأعضاء ناديين B , A (الوزن بالكيلو):

النادي A	32	28	41	39	34	
النادي B	36	39	30	34	39	35

اختبر فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ مع العلم بأن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٥٩- تعتقد الدراسة الحديثة أن خريجي المعاهد المتوسطة يتزوجون في عمر أقل من خريجي الجامعات. لتدعيم هذا الاعتقاد ، اختيرت عينة عشوائية من 100 فرد من كل مجموعة وتم تسجيل العمر عند الزواج وكانت النتائج كما يلي :

خريجي المعاهد المتوسطة $\bar{x}_1 = 22.5, \quad s_1 = 1.4$

خريجي الجامعات $\bar{x}_2 = 28, \quad s_2 = 1.5$

اختبر فرض العدم بعدم وجود فرق معنوي بين المجموعتين وذلك عند مستوى معنوية $\alpha = 0.01$.

-٦٠- تم تقسيم مجموعة من الأطفال حديثي الولادة في مستشفى إلى مجموعتين ، كل مجموعة استخدمت نوع من لبن الأطفال وقد تم تسجيل وزن الأطفال في كل مجموعة وذلك بعد 6 أسابيع من الولادة ، كانت النتائج كما يلي بالكيلو :

المجموعة A	3.0	4.2	4.5	5.0	5.2	4.6	6.1	5.6
المجموعة B	4.2	4.5	4.4	5.5	5.8	8.7	8.6	

المطلوب معرفة هل هناك فرق معنوي بين النوعين عند مستوى معنوية $\alpha = 0.1$ علما بأن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٦١- لمقارنة إحدى العناصر المعدنية لنوعين من العصائر B , A أخذت عينة عشوائية من العلب المعروضة في الأسواق لكل منهما وكانت النتائج كما يلي (وحدة القياس mg/100 gm)

النوع A	4.6	4.2	4.5	4.3	4.9	5.2	5.7
النوع B	5.1	4.3	4.2	5.7	5.8	5.1	5.3

أ- أختبر فرض العدم $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$.

ب- أختبر فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.02$ وذلك تحت فرض أن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٦٢- أعطي اختبار في القراءة لعينة من 12 طفل من أصل أمريكي وعينة أخرى من 10 أطفال من أصل مكسيكي . نتائج هذا الاختبار كانت :

أمريكي الأصل $\bar{x}_1 = 74 \quad , \quad s_1 = 8$

مكسيكي $\bar{x}_2 = 70 \quad , \quad s_2 = 10$

أ- أختبر فرض العدم $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$.

ب- أختبر فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$ وذلك تحت فرض أن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٦٣- لدراسة تأثير نوع من السماد على مساحة الورقة لنبات ما. اختير 16 قطعةً للزراعة وتم اختيار 8 منهم عشوائيا للمعالجة بالسماد والباقي تركت لمعالجة المراقبة وتم زراعة النبات في القطع التي عددها 16. البيانات التالية تعطى مساحة الورقة لكل نبات .

المعالجة بالسماد	1024	1216	1312	1780	
	1216	1312	992	1120	
معالجة المراقبة	1104	1072	1088	1328	1376
	1280	1120	1200		

أختبر فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ علما بأن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٦٤- تعطى النتائج التالية الزيادة في الوزن (gm) لعينتين الأولي عينة مراقبة والأخرى تم معالجة أفرادها بمركب Soft steroid (الجرعة 1 mg / pellet dose)

	n_i	$\bar{x}_i$	s_i
المراقبـــة	10	40.5	2.5
Soft Steroid	8	32.8	2.6

اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مســـتوى معنوية $\alpha = 0.1$ علما بأن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٦٥- في دراسة لأهمية دخول الطالب المعمل في مادة الفيزياء اختيرت عينة عشوائية مـــن 11 طالبا تم دخولهم المعمل. وكان متوسط الدرجة النهائية $\bar{x}_1 = 85$ بانحراف معياري $s_1 = 4.6$ كما اختبرت عينة أخرى من 17 طالبا من الذين لم يدخلوا المعمل فكان $\bar{x}_2 = 79$ بانحراف معيـــاري $s_2 = 6.2$. اختبر الفرض القائل أن دخول المعمل يؤثر على درجات الطلبة عند مستوى معنويـــة $\alpha = 0.1$ وذلك تحت فرض أن العينتين تم اختيارهما من توزعين طبيعيين .

-٦٦- إذا كانت نسبة الألياف في عينتين عشوائيين من شجيرات الكتان هي :

العينة الأولي	11.9	11.8	12.4	12.6	11.9	11.2
العينة الثانية	12.9	12.6	12.7	12.8	12.7	

المطلوب اختبار ما إذا كانت العينتين مأخوذتان من مجتمعين لهمـــا نفـــس المتوســـط والانحراف المعياري عند مستوى معنوية $\alpha = 0.1$ وذلك تحت فرض أن العينتين تم اختيارهما من مجتمعين طبيعيين .

-٦٧- البيانات التالية تمثل أزمنة العرض بالدقيقة لأفلام منتجة من شركتين A , B.

الشركة A	102	88	98	109	92	
الشركة B	81	166	98	93	88	110

اختبر الفرض القائل أن متوسط زمن العرض للفيلم المنتج من الشركة A يزيد عـــن متوســـط زمن العرض للفيلم المنتج من الشركة B وذلك تحت فرض أن توزيع الزمن للشركتين طبيعي (مستوى المعنوية $\alpha = 0.1$) .

-٦٨- يعتقد مدير شركة ما أن الأجر في الساعة بالدولار للعمال النصف ماهرة متســـاوية في قسمين من الشركة كالأتي :

القسم الأول : $\bar{x}_1 = 5.3$, $n_1 = 50$ $s_1 = 2.1$

القسم الثاني : $\bar{x}_2 = 5.9$, $n_2 = 40$ $s_2 = 1.7$

اختبر فرض العدم $H_0: \mu_1 = \mu_2$ عند مستوى معنوية $\alpha = 0.1$.

-٦٩- في مقارنة لمتوسط درجات الذكاء IQ لفرقتين في كلية ما اختيرت عينة عشـوائية مـن أربعة طلبة من إحدى الفرق وكانت درجات الذكاء لهم 117 ,116 ,113 ,110 . كمـا اختيرت عينة عشوائية من الفرقة الأخرى وكانت درجات الذكاء لهم 109 ,111 ,112 ,112 اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 > \mu_2$ عند مسـتوى معنويـة $\alpha = 0.1$ تحت فرض أن العينتين تم اختيارها من توزيعين طبيعيين .

-٧٠- اختيرت عينة عشوائية من البحارة حجمها $n = 10$ وقيست أوزانهم بالبوصة فكانت

80, 70, 68, 66, 72, 73, 70, 61, 60, 72

كما اختيرت عينة عشوائية أخرى من الجنود فكانت أوزانهم كمايلي :

61, 61, 72, 70, 73, 62, 63, 71, 70

أ- أختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مسـتوى معنوية $\alpha = 0.02$.

ب- أختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.01$ تحت فرض أن العينتين تم اختيارهما من توزيعين طبيعيين .

-٧١- تم قياس كمية الأنسولين لعينة من الفئران المصابة بمرض السكر والتي تعـالج بجرعـات منخفضة من الأنسولين (المجموعة A) وأخرى تعالج بجرعات عالية من الأنسولين (المجموعـة B) وتم الحصول على البيانات التالية :

المجموعة A $\quad n_1 = 8, \quad \bar{x}_1 = 1.98, \quad s_1 = 0.5,$

المجموعة B $\quad n_2 = 12, \quad \bar{x}_2 = 1.3, \quad s_2 = 0.35$

اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مسـتوى معنويـة $\alpha = 0.01$. تحت فرض أن العينتين تم اختيارهما من توزيعين طبيعيين .

-٧٢- البيانات التالية تمثل عدد البكتريا الهوائية (عدد المستعمرات/ قدم 3) والمأخوذة مـن 8 عينات مأخوذة من حجرات مفروشة بالسجاد و 8 عينات غير مفروشة بالسجاد في مستشفى ما

حجرات مفروشة	11.2	8.2	7.1	13	10.8	10.1	14.6	14
حجرات غير مفروشة	12.1	8.2	3.8	7.2	12	11.1	10.1	13.7

أختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مسـتوى معنوية فرض أن $\alpha = 0.02$ تحت العينتين تم اختيارهما من توزيعين طبيعيين .

-٧٣- إذا كانت عدد البكتريا بالملايين على نبات محفوظ على درجتي 10^o , 20^o بعـد 10 أيام هو :

$10^o \quad 8 \quad 9 \quad 6 \quad 7 \quad 7 \quad 8 \quad 9 \quad 10$

20   6   5   3   2   1   1   3   4

أ- أختبر فرض العدم $H_0 : \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1 : \sigma_1^2 \neq \sigma_2^2$ عنـــد مســـتوى معنوية $\alpha = 0.02$.

ب- أختبر فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$ تحت فرض أن العينتين تم اختيارهما من توزيعين طبيعيين .

-٧٤- طبق اختبار للقدرة على التفكير الناقد على مجموعة من المراهقين قبل حضورهم برنـــامج أعد لهذا الغرض مدته 40 أسبوعا وبعد حضورهم للبرنامج فإذا كان حجم العينة 10 أفـــراد تحقق من صحة الفرض القائل أن للبرنامج فاعلية على تنمية الفكر الناقد لـــدى المراهقـــين إذا كانت البيانات كايلي :

الشخص	1	2	3	4	5	6	7	8	9	10
قبل البرنامج	15	15	11	10	10	15	17	16	11	12
بعد البرنامج	20	25	14	16	22	23	24	25	24	26

وذلك عند مستوى معنوية $\alpha = 0.05$.

-٧٥- لمقارنة عليقتين من ناحية تأثيرهما على نمو العجول خلال شهرين من التغذية اختـــيرت عينة عشوائية من 10 أزواج (توائم) من العجول . اختبر صحة الفرض القائل $H_0 : \mu_D = 0$ ضد الفرض البديل $H_1 : \mu_D = 0$ عند مستوى معنوية $\alpha = 0.01$. إذا توافرت لك البيانـــات التالية :

التوأم	1	2	3	4	5	6	7	8	9	10
العليقة أ	20.1	30.6	31.7	27.2	22.7	22.4	29.3	28.8	26.3	27.1
العليقة ب	25.5	32.2	30.1	29.3	25.1	30.1	30.6	28.1	30.2	30.1

-٧٦- لدراسة الآثار الجانبية لدواء معين على ضغط الدم أخذت قياسات ضغط الـــدم لعينـــة عشوائية من المرض حجمها 8 قبل وبعد تناول الدواء فكانت البيانات كالتالي :

الشخص	1	2	3	4	5	6	7	8
قبل الدواء	70	72	80	65	66	77	70	72
بعد الدواء	66	70	75	62	64	73	74	77

اختبر الفرض القائل $H_0 : \mu_D = 0$ ضد الفرض البديل $H_1 : \mu_D \neq 0$ عند مستوى معنويـــة $\alpha = 0.05$.

-٧٧- طبق مقياس للاتجاهات نحو الأطفال في عينة من السيدات الآتي تم قبولهم بقسم دراسات الطفولة فور التحاقهن بالقسم ، ثم طبق نفس المقياس مرة أخرى فور حصولهن على البكالوريوس وكانت النتائج كالتالي :

السيدة	1	2	3	4	5	6	7	8	9	10
عند الالتحاق	9	6	6	14	14	8	8	12	8	14
عند التخرج	11	20	15	15	15	18	18	17	16	16

اختبر الفرض العدم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

-٧٨- قام شخص بإجراء 6 عمليات حسابية على آلتين حاسبتين وتسجيل الزمن اللازم لكل عملية على كل آلة في الجدول التالي :

العملية الحسابية	1	2	3	4	5	6
الآلة الأولى	22	17	28	21	32	19
الآلة الثانية	18	19	23	22	30	22

اختبر الفرض العدم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

-٧٩- للمقارنة بين طريقتين لتقدير كمية اللبن المهضوم في الأطفال الرضع ، اختيرت عينة عشوائية من الحجم $n = 14$ ، البيانات التالية تعطى كمية اللبن المهضوم لكل طفل في العينة :

الطفل	1	2	3	4	5	6	7
Isotopic	1509	1418	1561	1556	2169	1760	1098
Test-weighing	1448	1254	1336	1565	2000	1318	1410
الطفل	8	9	10	11	12	13	14
Isotopic	1198	1479	1781	1414	1954	2174	2058
Test-weighing	1179	1342	1124	1468	1604	1722	1518

اختبر الفرض العدم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

-٨٠- للمقارنة من نوعين من مصايد الأسماك الخاصة بسمك الهامور ، استخدم النوعين خلال 16 فترة زمنية خلال 4 سنوات وذلك في بحيرة ما و البيانات التالية تعطى كمية الأسماك لكل يوم

الفترة	1	2	3	4	5	6	7	8
Pipe	6.64	7.89	1.83	0.42	0.85	0.29	0.57	0.63
Brush	9.73	8.1	2.17	0.75	1.61	0.75	0.83	0.56
الفترة	9	10	11	12	13	14	15	16
Pipe	0.32	0.37	0.0	0.11	4.86	1.8	0.32	0.88
Bursh	0.67	0.32	0.98	0.52	5.38	2.33	0.91	0.70

اختبر الفرض العدم $H_0 : \mu_D = 0$ ضد الفرض البديل $H_1 : \mu_D \neq 0$.

٨١- من المعروف أن العناصر النادرة لها تأثير غير مرغوب فيه على طعم مياه الشرب وأيضـا آثار ضارة على الصحة إذا وجدت بكميات كبيرة . قام باحث بإخبار الزنك (مقـاس mg/L) لكل من سطح وقاع النهر في كل موقع . الأزواج الستة من المشاهدات معطـاة في الجدول التالي . هل تدل هذه البيانات على أن المتوسط الحقيقي لتركيز الزنك في قاع الماء يزيـد عن سطح الماء ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الموقع	1	2	3	4	5	6
قاع النهر	0.430	0.266	0.567	0.531	0.707	0.716
سطح النهر	0.415	0.238	0.390	0.410	0.605	0.609

٨٢- لاختبار تأثير إضافة مواد الجفاف على الدهان أجريت تجربة على 6 عينات من المعـدن حيث قسمت كل عينة إلى نصفين لكل عينة ثم دهان نصف العينة بالدهان المضـاف لـه مـواد الجفاف بينما تم دهان النصف الآخر من العينة بالدهان بدون إضافة مـواد الجفـاف . وتـترك العينات التي عددها 12 حتى الجفاف وتم تسجيل زمن الجفاف في الجدول التالي :

العينـة	1	2	3	4	5	6
الدهان مضـاف له مواد الجفاف	3.4	3.8	4.2	4.1	35	4.7
الدهان بـدون إضافة	3.6	3.8	4.3	4.3	3.6	4.6

اختبر الفرض العدم $H_0 : \mu_D = 0$ ضد الفرض البديل $H_1 : \mu_D \neq 0$ وذلك عند مسـتوى معنوية $\alpha = 0.05$

-٨٣- البيانات التالية تمثل أوزان 10 قوائم عند لحظة الميلاد (مقاس بالكيلو) وذلك عند مستوى معنوية $\alpha = 0.05$

التوائم	1	2	3	4	5	6	7	8	9	10
الأول في الولادة	3.95	3.41	3.73	3.48	4.28	3.98	4.18	4.04	3.73	4.13
الثاني في الولادة	3.93	3.35	3.72	4.18	3.44	4.15	3.89	4.20	4.00	3.72

اختبر فرض العدم أن الطفل الأول في النزول لأي توائم أثقل في الوزن من الطفل الثاني .

-٨٤- إذا كان من المعلوم أن واحد من كل سبعة يدخنون في إحدى الدول . فإذا أجريت حملة للتوعية عن مضار التدخين . للحكم على مدي نجاح تلك الحملة أخذت عينة عشوائية من 600 شخص ووجد من بينهم 180 لا يزالون يدخنون هل البيانات تعطى دليلا كافيا على نجاح الحملة وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨٥- أصيبت إحدى المناطق الزراعية التي مساحتها 6000 فدان بإحدى الآفات الزراعية . اختيرت عينة عشوائية حجمها 300 فدان فوجد أن نسبة الإصابة بها %20 اختبر الفرض القائل أن نسبة الإصابة %15 عند مستوى معنوية $\alpha = 0.01$.

-٨٦-أجريت دراسة على المركز المالي لعينة عشوائية حجمها 200 من شركات الغزل والنسيج فوجد أن %25 منها سوف تتجه إلى الإفلاس بعد عام . اختبر الفرض القائل أن %40 من شركات الغزل والنسيج في المجتمع المسحوب منه العينة سوف تتجه إلى الإفلاس بعد عام وذلك عند مستوى معنوية $\alpha = 0.01$.

-٨٧- اختيرت عينة عشوائية من 200 شخص من مجتمع ما ووجد أن 40 شخص من العينة مصابون بمرض ما . المطلوب اختبار الفرض أن نسبة الإصابة بالمرض في هذا المجتمع أقل من 0.5 وذلك عند مستوى معنوية $\alpha = 0.01$.

-٨٨- يعتقد مدير الإنتاج في مصنع لإنتاج التلفزيونات في بلد ما أن %80 من الأسر تمتلك تلفزيون ملون . للتحقق من هذا الفرض اختيرت عينة عشوائية من 1000 أسرة ووجد أن 318 منهم يمتلكون تليفزيونا ملونا اختبر صحة هذا الفرض $(p = 0.8)$ عند مستوى معنوية $\alpha = 0.05$.

-٨٩- يعتقد المسئولين في كلية ما أن 25% من الطلبة يمتلكون سيارة . اختبر صحة هذا الفرض إذا اختيرت عينة عشوائية من 90 طالبا ووجد أن 28 منهم يمتلك سيارة وذلك عند مستوى معنوية $\alpha = 0.05$.

-٩٠- في دراسة لتقدير نسبة ربات البيوت اللاتي يمتلكن غسالة بمجفف في بلدنا ما ، اختيرت عينة عشوائية من 1000 ربة منزل ووجد أن 630 منهم يمتلكن غسالة بمجفف . اختبر صحة الفرض القائل أن $p = 0.4$ ضد الفرض البديل $p \neq .4$ عند مستوى معنوية $\alpha = 0.05$.

-٩١- في تقرير خاص وجد أن 30% فقط من خريجي الجامعات يجدون عملا في مجال تخصصهم الرئيسي اختيرت عينة عشوائية من 500 خريج فوجد أن بينهم 250 يعملون في مجال تخصصهم اختبر فرض العدم $H_0: p = .3$ ضد الفرض البديل $H_1: p \neq .3$ عند مستوى معنوية $\alpha = 0.01$.

-٩٢- يعتقد أن نسبة الأسر التي تشترى اللبن من المصنع A في مدينة ما هو $p = 0.6$ فإذا اختيرت عينة عشوائية من 200 أسرة ووجد أن نسبة الأسر التي تشترى اللبن 80. اختبر فرض العدم $p = 0.6$ عند مستوى معنوية $\alpha = 0.05$.

-٩٣- في عينة عشوائية من 400 ناخب في مدينة ما وجد أن 200 منهم يوافقون على فرض 3% ضريبة مبيعات . اختبر فرض العدم أن $p = 0.45$ ضد الفرض البديل $p \neq 0.45$ عند مستوى معنوية $\alpha = 0.01$.

-٩٤- يعتقد أن 40% من الطلبة في كلية ما يستخدمون نظارة طبية اختيرت عينة عشوائية من 64 طالبا ووجد أن 40 منهم يستخدم نظارة طبية اختبر فرض العدم أن $p = 0.4$ ضد الفرض البديل $p \neq 0.4$ عند مستوى معنوية $\alpha = 0.01$.

-٩٥- تريد شركة للتليفونات أن تقرر فيما إذا كان بعض الخطوط الجديدة يمكن تجهيزها تحت الأرض. ولأن الرسوم المضافة إلى فاتورة التليفون قليلة بالنسبة إلى التكاليف الباهظة التي تتحملها الشركة لذلك قررت الشركة عمل استفتاء للعملاء وإذا كان 60% من العملاء يفضلون تركيب تحت الأرض فإنها سوف توافق على تجهيز الخطوط الجديدة تحت الأرض وغير ذلك ترفض فإذا كان 118 من العملاء وافق على التركيب من بين 160 عميل اختبر فرض العدم $H_0: p = 0.6$ ضد الفرض البديل $H_1: p \neq 0.6$ عند مستوى معنوية $\alpha = 0.05$.

-٩٦- في دراسة عن نسبة الكوابيس الليلية اختيرت عينة عشوائية من 160 رجلا وعينة عشوائية أخرى من 192 سيدة وتم سؤالهم على عدد الكوابيس التي تعرضوا لها (على الأقل

خلال شهر) فوجد أن 55 من الرجال و 60 سيدة تعرضوا للكوابيس اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.05$.

-٩٧- في عينة عشوائية من 300 شاب من المدينة A وجد أن 63 منهم يفضلون قيادة سيارتهم في الطريق الصحراوي على سرعة من 55 , 65 mph بينما وجد أن 75 من 180 في المدينة B يفضلون السرعة من 55 إلى 65 mph . اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.05$.

-٩٨- في عينة عشوائية من 5726 رقم تليفون في منطقة ما في مارس سنة 1992 وجد أن 1105 غير مقيدين في الدليل وبعد سنة ومن عينة عشوائية من 5384 وجد أن 1001 غير مقيدين في الدليل . اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.05$.

-٩٩- في دراسة عن نسبة الذين يفضلون استخدام حزام الأمان أخذت عينة عشوائية من 200 سائق يعمل على سيارة الغير ووجد أن 115 منهم يفضل استخدام حزام الأمان بينما في عينة أخرى من 300 شخص (يقود سيارته الخاصة) وجد أن 154 يفضل استخدام حزام الأمان . اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.05$.

-١٠٠- ترغب شركة في اختبار ما إذا كانت نسبة المقبول من المكونات الإلكترونية لمورد أجنبي p_1 تزيد عنها لمورد محلي p_2 . اختيرت عينة عشوائية من شحنه كل مورد وتم الحصول على البيانات التالية :

$$n_1 = 200 \quad , \quad n_2 = 200 \quad , \quad \hat{p}_1 = 0.77 \quad , \quad \hat{p}_2 = 0.8$$

اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.01$.

-١٠١- قام أحد العاملين في مكتبة بعمل إحصائية عن مدى تفضيل المترددين للكتب الثقافية . اختيرت عينة عشوائية من السيدات المترددين على المكتبة فوجد أن 400 من 600 يفضلون قراءة الكتب الثقافية ، كما اختيرت عينة عشوائية من الرجال المترددين على المكتبة فوجد أن 500 من 700 يفضلون قراءة الكتب الثقافية، اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.01$.

-١٠٢- في دراسة ميدانية لمجموعتان من الأشخاص عن رأيهم في برنامج تليفزيوني المجموعة الأولي مكونة من 600 شخص من الريف بينهم 300 يفضلون البرنامج المجموعة الثانية والمكونة

من 1000 شخص من المدن بينهم 600 شخص يفضلون البرنامج . هل البرنامج شائع في الريف والمدينة بنفس النسبة وذلك عند مستوى معنوية $\alpha = 0.05$.

-١٠٣- من مجموعتين متشابهتين ومصابتين بمرض ما ، اختيرت من المجموعة الأولى 50 شخصا ومن المجموعة الثانية 80 شخصا وقد استخدم مع كل عينة نوع مختلف من الدواء فإذا كان عدد الذين تم شفائهم من العينة الأولى 30 وعدد من شفى من المجموعة الثانية 40 فهل هناك دلائل تشير إلى اختلاف نسبة الشفاء باستخدام كل دواء عند مستوى معنوية $\alpha = 0.05$.

-١٠٤- في دراسة عن نسبة ربات البيوت اللاتي يمتلكن غسالة بمجفف في منطقة بين A , B اختيرت عينة عشوائية من الحجم 1000 من المنطقة A فكان عدد اللاتي يمتلكن غسالة 630 . كما اختيرت عينة عشوائية أخرى من المنطقة B حجمها 1200 ووجد أن 710 منهن يمتلكن غسالة بمجفف. اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1 : p \neq p_2$ عند مستوى معنوية $\alpha = 0.01$.

الفصل العاشر

الانحدار والارتباط

Regression and Correlation

(١٠-١) الانحدار Regression

يستخدم الانحدار لدراسة العلاقات بين متغيرات قابلة للقياس. تحليل الانحدار له تطبيقات كثيرة في معظم مجالات الدراسة التي تشتمل على العلوم الحيوية ، الفيزيائية ، العلوم الاجتماعية ، الاقتصادية ، الصناعية ... الخ. تختلف الطرق المستخدمة في تحليل الانحدار باختلاف التطبيقات ولكنها تشترك جميعها في الوصف والتنبأ بالقيم المستقبلية . عموما يتكون الانحدار من عدة خطوات. فللدراسة العلاقة بين عدد من المتغيرات ، تجمع البيانات على عينة من المفردات المعرضة لتلك المتغيرات. في نموذج الانحدار يكون هناك متغير واحد يسمي المتغير التابع depend variable أو متغير الاستجابة response variable ، بينما المتغيرات الأخرى تسمي متغيرات مستقلة independent variables . تستخدم البيانات للحصول على تقديرات لمعالم النموذج . في الحقيقة طرق تقدير معالم النموذج المجهولة كثيرة . سوف تقتصر دراستنا على طريقة المربعات الصغرى least squares. يشتمل البند التالي على الانحدار الخطى البسيط ، أي وجود متغير تابع ومتغير مستقل واحد والعلاقة بينهما خطية ، حيث ينحصر اهتمامنا على تعريف النموذج المناسب ومناقشة الفروض وإيجاد تقديرات المربعات الصغرى وعرض بعض طرق التقدير بفترة واختبارات الفروض .

(١٠-٢) الانحدار الخطى البسيط Simple Linear Regression

بفرض أننا اخترنا عينة عشوائية من الحجم n من المجتمع موضع الدراسة ممثلة بالفئة $\{(x_i , y_i) \; ; i = 1,2,...,n\}$. بتكرار المعاينة مع استخدام نفس قيم x ، فإننا نتوقع أن قيم y تختلف من عينه إلى أخرى. وعلى ذلك فإن قيمة y_i في الزوج المرتبة (x_i , y_i) هي قيمة لمتغير عشوائي Y_i .سوف نرمز للتوزيع الاحتمالي للمتغير $Y \mid x$ المقابل لقيمة ثابتة x بالرمز $f(y \mid x)$. من الواضح ، أنه إذا كانت $x = x_i$ ، فإن الصيغة $Y \mid x_i$ تمثل متغير عشوائي Y_i . اهتمامنا سوف يكون في توزيعات فئة المتغيرات العشوائية $\{Y_i , i = 1,2,...,n\}$. والتي يفترض أنها مستقلة. أيضا للحصول على فترات ثقة واختبارات فروض ، لابد أن تكون $Y_1, Y_2,..., Y_n$ تتبع توزيعا طبيعيا .

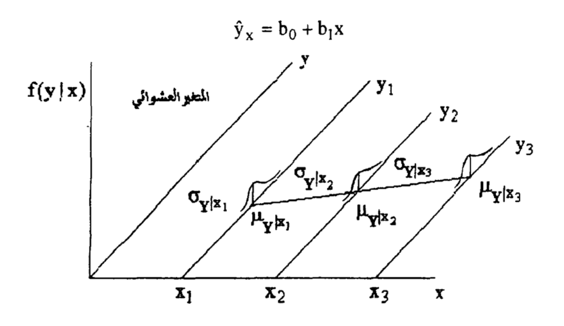
في مشكلة الانحدار سوف نعرف متوسط توزيع Y عند قيمة معطاة x بالرمز $\mu_{Y|x} = E(Y \mid x)$ وتباين توزيع Y بالرمز $\sigma^2_{Y|x}$ وذلك تحت فرض أن تباينات المتغيرات $Y_1, Y_2,..., Y_n$ متساوية ، أي أن $\sigma^2_{Y|x} = \sigma^2$ ، لجميع قيم x و المعلمة $\mu_{Y|x}$ ثابتة لأي x ولكن قد تختلف باختلاف قيم x .

المنحنى الذي يربط متوسط التوزيعات للمتغيرات $Y_1, Y_2, ..., Y_n$ يسمى خط الانحـــدار regression curve . إذا وقعت جميع المتوسطات $\mu_{Y|x}$ على خط مستقيم كما هو موضح في شكل (١٠-١) فإن الانحدار يكون خطي ويمكن تمثيله بالمعادلة :

$$\mu_{Y|x} = \beta_0 + \beta_1 \ x,$$

حيث β_1 , β_0 معالم النموذج المجهولة والتي يراد تقديرها من بيانات العينـــة. ســـوف نرمـــز لتقديرات هذه المعالم بالرموز b_0 , b_1 على التوالي . وعلى ذلك يمكن تقدير $\mu_{Y|x}$ بواســـطة $\hat{y}_x$ من خط الانحدار المقدر التالي :

$$\hat{y}_x = b_0 + b_1 x$$



شكل (١٠-١)

(١٠-٢-١) شكـــل الانتشار Scatter Plot

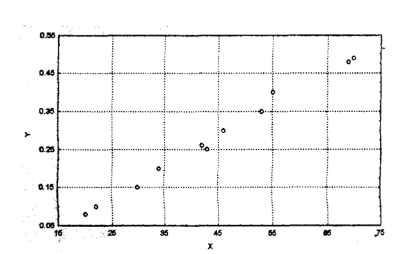
الأسلوب المفيد لبدء تحليل الانحدار البسيط هو تمثيل البيانات بيانياً وهو ما يعرف بشــــكل الانتشار scatter plot وذلك لمحاولة اكتشاف الصورة التقريبية للعلاقة. للحصول على شكل الانتشار يخصص محور x (المحور الأفقي) للمتغير المستقل بينما يخصص محور y (المحور الرأســـي) للمتغير التابع . لكل زوج (x , y) من أزواج المشاهدات التي عددها n نقوم بتوقيع نقطـــة على الرسم. كثير من البرامج الإحصائية مثل برنامج Statistica SPSS يمكن اســـتخدامهم للحصول على أشكال الانتشار .

مثال (١٠-١) في إحدى التجارب وزن قرون عدد من الغزلان المختلفة الأعمار و كانت النتائج كما هي مدونة في الجدول (١٠-١) . المطلوب رسم شكل الانتشار وتحديد شكل العلاقة بين المتغيرين .

جدول (١٠-١)

العمر x	20	22	30	34	42	43	46	53	55	69	70
الوزن y	0.08	0.10	0.15	0.20	0.26	0.25	0.30	0.35	0.40	0.48	0.49

الحل . يتضح من شكل (١٠-٢) أن النقط عموماً ، ليس بالضبط ، تقع على خط مستقيم. هذا يجعلنا نقترح أن العلاقة بين المتغيرين يمكن وصفها (كتقريب أولي) بمعادلة خط مستقيم .



شكل (١٠-٢)

(١٠-٢-٢) بناء نموذج الانحدار البسيط Building a Simple Regression Line

بفرض أن كل المتوسطات ، $\mu_{Y|x}$ تقع على خط مستقيم وعلى ذلك يمكن كتابة خط انحدار المجتمع على الشكل :

$$\mu_{Y|x} = \beta_0 + \beta_1 x$$

كما يمكن كتابة المتغير العشوائي $Y_i = Y \mid x_i$ على الشكل :

$$Y_i = \beta_0 + \beta_1 x_i + E_i ,$$

حيث E_i متغير عشوائي ، بناء على الفروض السابقة على Y_i ، من الضروري أن يكون له متوسط صفر وتباين $\sigma^2_{Y|x} = \sigma^2$. كل مشاهدة $(x_i \ , \ y_i)$ في العينة لا بد أن تحقق العلاقة :

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i .$$

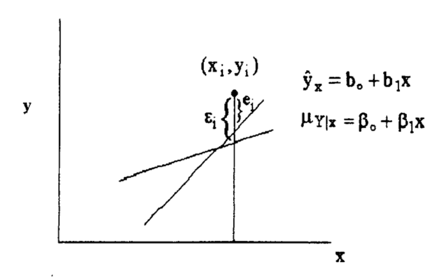
لعينة عشوائية من الحجم n من المشاهدات $\{(x_i, y_i), i = 1,2,...,n\}$ فإن خط الانحـــدار ، الذي يعتبر تقدير لـــ $\mu_{Y|x}$، يمكن إيجاده بالمعادلة :

$$\hat{y}_x = b_0 + b_1 x,$$

وكل زوج من المشاهدات لا بد أن يحقق العلاقة :

$$y_i = b_0 + b_1 x_i + e_i$$

حيث e_i يسمى الباقي residual. الفرق بين ϵ_i و e_i موضح في شكل (١٠-٣).



شكل (١٠-٣)

يعرف مجموع مربعات البواقي كالآتي :

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} [y_i - b_0 - b_1 x_i]^2$$

ويرمز له بالرمز S S E والذي غالباً ما يسمى مجموع مربعـــات الأخطـــاء sum of squares of the errors حول خط الانحدار .

(١٠-٢-٣) طريقة المربعات الصغرى Method of Least Squares

سوف نوجد التقديرين b_0 , b_1 للمعالم $\beta_0 \;,\; \beta_1$ على التـــوالي بحيـــث أن مجمـــوع مربعات الأخطاء يكون أقل ما يمكن. هذه الطريقة لتقدير المعالم تسمى طريقة المربعات الصغرى.

باستخدام حساب التفاضل نجد أن قيمتي β_1 , β_0 اللتين تحققان النهاية الصغـرى لمجمـوع مربعات الأخطاء يمكن تقديرهما بقيمتي b_1 , b_0 اللتين تحققان المعادلتين :

$$b_0\ n + b_1 \Sigma\ x_i = \Sigma\ y_i,$$

$$b_0\ \Sigma\ x_i + b_1 \Sigma\ x_i^2 = \Sigma\ x_i\ y_i$$

وتسميان بالمعادلتين الطبيعيتين وبحل هاتين المعادلتين آنيا نحصل على :

$$b_1 = SXY / SXX,$$

$$SXX = \sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n},$$

$$SXY = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n},$$

$$b_0 = \bar{y} - b_1\ \bar{x}.$$

مثال (١٠-٢) أجريت تجربة لدراسة العلاقة بين التسميد ومحصول الذرة . البيانات الـتي تم الحصول عليها معطاة في جدول (١٠-٢)

جدول (١٠-٢)

x السماد	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4
y المحصول	10	15	30	35	25	30	50	45

(أ) أرسـم شكل الانتشار

(ب) أوجد معادلة خط الانحدار المقدرة

الحل . (أ) يتضح من شكل الانتشار (١٠-٤) أن الخط المستقيم هو أفضل طريقة لتمثيل هذه البيانات :

أي أننا نفترض النموذج الخطى البسيط :

(ب) بما أن β_1 , β_0 مجهولتان فإننا نقدرهما من مشاهدات العينة حيث :

$$n = 8 \quad \Sigma\ x_i = 10.8 \quad \Sigma\ x_i^2 = 18.36$$

$$\Sigma\ x_i y_i = 385.5 \quad , \quad \bar{x} = 1.35 \quad , \quad \bar{y} = 30, \quad \Sigma\ y_i = 240.$$

$$b_1 = \frac{SXY}{SXX} = \frac{\Sigma x_i y_i - \frac{\Sigma x_i \Sigma y_i}{n}}{\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}}$$

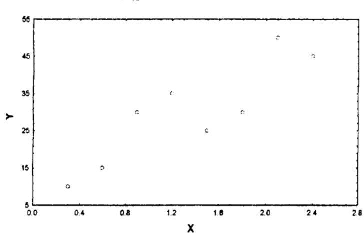
$$= \frac{385.5 - \frac{(10.8)(240)}{8}}{18.36 - \frac{(10.8)^2}{8}}$$

$$= \frac{61.5}{3.78} = 16.27,$$

$$b_0 = \bar{y} - b_1\bar{x} = 30 - (16.27)(1.35) = 8.036.$$

معادلة الانحدار المقدرة سوف تكون على الشكل :

$$\hat{y}_x = 8.036 + 16.27 \ \ x.$$



شكل (١٠-٤)

(١٠-٢-٤) تحليل الانحدار **Analysis of Variance**

لاختبار معنوية معامل الانحدار β_1 أي اختبار فرض العدم :

$$H_0 : \beta_1 = 0$$

ضد الفرض البديل :

$$H_1 : \beta_1 \neq 0$$

يجب دراسة مكونات مجموع المربعات وتجزئته إلى مكوناته الأساسية على النحو التالي :

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$$

وذلك بإضافة وطرح $\hat{y}_i$ ($\hat{y}_i = b_0 + b_1 x_i$) من الطرف الأيمن وتربيـــع طـــرفي المعادلـــة والتجميع مع ملاحظة أنه يمكن إثبات أن المقدار التالي مساوي الصفر :

$$\Sigma(\hat{y}_i - \bar{y})(y_i - \hat{y}_i)^2 = 0.$$

وعلى ذلك :

$$\Sigma(y_i - \bar{y})^2 = \Sigma(\hat{y}_i - \bar{y})^2 + (y_i - \hat{y}_i)^2.$$

أي أن مجموع المربعات حول المتوسط يساوى مجموع مربعات الانحـــدار مضـــاف إلى مجمـــوع المربعات حول الانحدار ويمكن الرمز إليه بالمتساوية التالية :

$$\mathbf{SSTO = SSR + SSE}.$$

وهناك صيغ مبسطة للقيم **SSTO , SSR , SSE** حيث :

$$SSTO = SYY = \Sigma y_i{}^2 - \frac{\Sigma y_i^2}{n},$$

$$SSR = \frac{(SXY)^2}{SXX},$$

$$\mathbf{SSE = SSTO - SSR}.$$

من الناحية الإحصائية نجد أن لكل مجموع مربعات درجات حرية خاصة به ، فإذا كان لدينا n من المشاهدات فإن توزيع درجات الحرية يكون على الشكل الموضح في جدول (١٠-٣)

جدول (١٠-٣)

مجموع المربعات	درجات الحرية
مجموع مربعات الانحدار	1
مجموع مربعات الخطـــأ	n-2
مجموع المربعات الكلـــي	n − 1

بقسمة مجموع المربعات بدرجات الحرية الخاصة به نحصل على ما يســـمي متوســـط المربعـــات **mean squares** ويعتبر تباين العينة s^2 مثال لمتوسط المربعات. وعلى ذلك متوسط مجمـــوع مربعات الانحدار نرمز له بالرمز **MSR** ، هو :

$$MSR = \frac{SSR}{1}.$$

ومتوسط مربعات الخطأ ، نرمز له بالرمز **MSE** ، هو :

$$MSE = \frac{SSE}{n-2}.$$

من النتائج السابقة يمكن اشتقاق جدول تحليل التباين ANALYSIS OF VARIANCE ، للاختصار جدول ANOVA ، والموضح في جدول (١٠-٤) . الآن :

لاختبار فرض العدم $H_0 : \beta_1 = 0$ ضد الفرض البديل $H_1 : \beta_1 \neq 0$ وباعتبار أن فرض العدم صحيح فإن :

$$f = \frac{MSR}{MSE},$$

قيمة لمتغير عشوائي F يتبع توزيع F بدرجات حرية $v_1 = 1, \quad v_2 = n-2$. لمستوى معنوية α منطقة الرفض $F > f_\alpha(1, n-2)$ حيث $f_\alpha(1, n-2)$ تستخرج من جدول توزيع F في ملحق (٦) أو ملحق (٧) بدرجات حرية $v_1 = 1, \quad v_2 = n-2$. إذا وقعت f في منطقة الرفض نرفض H_0.

جدول (١٠-٤)

الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات
الانحدار	1	SSR	$MSR = \frac{SSR}{1}$
الخطأ	n-2	SSE	$MSE = \frac{SSE}{n-2}$
الكلي	n − 1	SSTO	

مثال (١٠-٣) لأزواج القياسات المعطاة في جدول (١٠-٥) :

المطلوب : (أ) إيجاد معادلة الانحدار الخطي المقدرة .

(ب) اختبار فرض العدم $H_0 : \beta_1 = 0$ ضد الفرض البديل $H_1 : \beta_1 \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

جدول (١٠-٥)

x	4	6	2	5	7	6	3	8	5	3	1	5
y	197	272	100	228	327	279	148	377	238	142	66	239

الحل .

$$n = 12 \quad \Sigma\, x_i = 55 \quad \Sigma\, x_i y_i = 14060$$

$$\Sigma\, y_i = 2613 \quad , \quad \bar{x} = 4.58333 \quad , \quad \bar{y} = 217.75,$$

$$\Sigma x_i^2 = 299 \ , \ \Sigma y_i^2 = 661865,$$

$$SXY = \Sigma x_i y_i - \frac{\Sigma x_i \Sigma y_i}{n}$$

$$= 14060 - \frac{(55)(2613)}{12} = 2083.75,$$

$$SXX = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

$$= 299 - \frac{(55)^2}{12} = 46.91667,$$

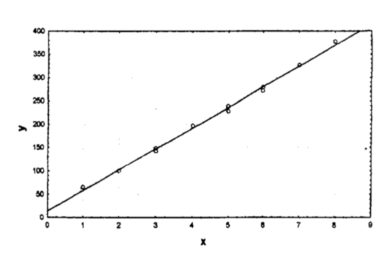
$$b_1 = \frac{SXY}{SXX} = \frac{2083.75}{46.91667} = 44.41385,$$

$$b_0 = \bar{y} - b_1 \bar{x}$$
$$= 217.75 - (44.41385)(4.58333)$$
$$= 14.187.$$

وعلى ذلك فإن معادلة خط الانحدار المقدرة هي :

$$\hat{y}_x = 14.187 + 44.41385.$$

والموضحة في شكل (١٠-٥) .



شكل (١٠-٥)

الآن نحسب :

$$SSTO = SYY = \Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}$$

$$- 661865 - \frac{(2613)^2}{12} = 92884.25,$$

$$SSR = \frac{(SXY)^2}{SXX} = \frac{(2083.75)^2}{46.91667}$$
$$= 92547.362.$$

وبطرح SSR من SSTO نحصل على :

$$SSE = SSTO - SSR$$
$$= 92884.25 - 92547.362$$
$$= 336.888,$$
$$MSR = \frac{SSR}{1} = \frac{92547.362}{1} = 92547.362.$$

$$MSE = \frac{SSE}{n-2} = \frac{336.888}{10} = 33.6888.$$

جدول تحليل التباين معطى في جدول (١٠-٦).

جدول (١٠-٦)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات
الانحدار	1	92547.362	92547.362
الخطأ	10	336.888	33.6888
الكلي	11	92884.25	

$$f = \frac{MSR}{MSE} = \frac{92547.362}{33.6888}$$
$$= 2747.126.$$

$f_{0.05}(1,10) = 4.96$ والمستخرجة من جدول توزيع F في ملحق (٦) بدرجات حريـــة $v_1 = 1, v_2 = 10$. منطقة الرفض $F > 4.96$. وبما ن f تقع في منطقة الرفض نرفض H_0 .

(١٠-٢-٥) تقدير σ^2 Estimating σ^2

في الحقيقة ، يفضل الحصول على تقدير للمعلمة σ^2 لا يعتمد على النمـــوذج. عمومـــا، مثل هذا التقدير يمكن الحصول عليها فقط في فئات البيانات التي تحتوي على قيم متعددة لـــ y عند كل قيمة من قيم x أو عند توفر بعض المعلومات القبلية. في حالة عدم توفر هذه الحـــالات فإن التقدير للمعلمة σ^2 يعتمد على النموذج ، أي يكون دالة في مجموع مربعات الأخطـــاء SSE ، ويحسب من المعادلة الآتية :

$$s^2 = \frac{SSE}{n-2} = MSE$$

أي أن s^2 يساوى متوسط مجموع مربعات الخطأ وهو تقدير غير متحيز للمعلمة σ^2. التقديـــر بنقطة للمعلمة σ هو $\sqrt{s^2}$ والذي يسمي الخطأ المعياري للانحدار **standard error of regression**. للمثال (١٠-٣) ومن جدول (١٠-٦) فإن :

$$s = \sqrt{MSE} = \sqrt{33.6888} = 5.804.$$

(١٠-٢-٦) معامل التحديد البسيط <u>Coefficient of Simple Determination</u>

يعرف معامل التحديد البسيط r^2 كالتالي :

$$r^2 = \frac{SSR}{SSTO} = \frac{SSTO - SSE}{SSTO} = 1 - \frac{SSE}{SSTO}$$

للمثال (١٠-٣) ومن جدول (١٠-٦) فإن :

$$r^2 = \frac{92547.362}{92884.25} = 0.996.$$

يأخذ r^2 الواحد الصحيح عندما تقع القيم $y_1, y_2, ..., y_n$ على خط الانحدار المقدر وعلـــى ذلك فإن $SSE = 0$. وذلك لأن :

$$SSR = SSTO - SSE$$

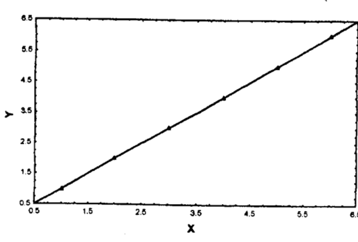
وفي هذه الحالة فإن :

$$SSR = SSTO - 0.0 = SSTO$$

وعلى ذلك فإن :

$$r^2 = \frac{SSR}{SSTO} = \frac{SSTO}{SSTO} = 1$$

هذه الحالة موضحة في شكل (١٠-٦) .

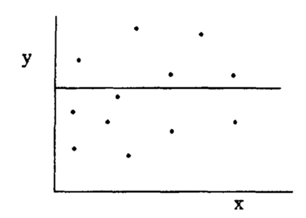


شكل (١٠-٦)

عندما $r^2 = 0$ فهذا يدل على عدم وجود علاقة خطيه بين المتغيرين وبالتالي فإن **SSR = 0** ومنها :

$$r^2 = \frac{SSR}{SSTO} = \frac{0.0}{SSTO} = 0.0$$

في هذه الحالة فإن معادلة الانحدار المقدرة سوف تكون موازية للمحور الأفقي ، أي أن $b_1 = 0$ كما هو موضح في شكل (١٠-٧)

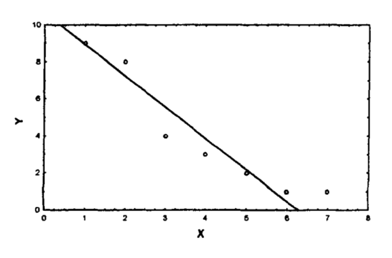


شكل (١٠-٧)

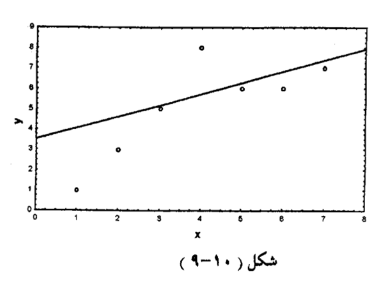
معامل التحديد دائما موجب وتتراوح قيمته بين الصفر والواحد الصحيح أي أن :

$$0 \le r^2 \le 1.$$

يتضح من شكل (١٠-٨) ، حيث $r^2 = - 0.9025$ ، أن المشاهدات تقترب بدرجة كبيرة من خط الانحدار المقدر وذلك بالمقارنة للمشاهدات في شكل (١٠-٩) حيث $r^2 = 0.3249$.



شكل (١٠-٨)



شكل (١٠-٩)

(١٠-٢-٧) تقدير المعلمتين β_0, β_1 Estimation of the parameters

في دراستنا السابقة في البند (١٠-٢) عرفنا أن b_0 , b_1 تقديرين للمعلمتين الحقيقتين β_0, β_1 وذلك بالاعتماد على عينة عشوائية من الحجم n. بتكرار المعاينة من الحجم n وحساب b_0 , b_1 لكل عينة فإن القيمتين b_0 , b_1 سوف يختلفان من عينة إلى أخرى . وعلى ذلك فإن التقديرين b_0 , b_1 قيمتين لمتغيرين عشوائية B_0, B_1 .

وبما أن قيمة x لازالت ثابتة ، فإن قيم β_0, β_1 سوف تعتمد على قيم y أو بدقة أكثر ، على قيم المتغيرات العشوائية $Y_1, Y_2, ..., Y_n$ المستقلة والتي تتبع توزيعاً طبيعياً ، وباستخدام نظريات الإحصاء يمكن إثبات أن المتغير العشوائي B_1 أيضاً يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{B_1} = \beta_1$$

وتباين :

$$\sigma_{B_1}^2 = \frac{\sigma^2}{SXX}.$$

وتبعاً لنظرية (٨-١) فإن :

$$Z = \frac{B_1 - \beta_1}{\sqrt{\frac{\sigma^2}{SXX}}}$$

متغير عشوائي يتبع التوزيع الطبيعي القياسي .

عادة الانحراف المعياري ، σ، مجهول ويستبدل بالإحصاء S في صيغة Z لنحصل على

$$T = \frac{B_1 - \beta_1}{\sqrt{\frac{S^2}{SXX}}}$$

حيث T متغير عشوائي له توزيع t بدرجات حرية n-2 . سوف نستخدم المتغير T في إيجاد $(1-\alpha)100\%$ فترة ثقة للمعلمة β_1 حيث :

$$P(-t_{\alpha/2} < \frac{B_1 - \beta_1}{\sqrt{S^2/SXX}} < t_{\alpha/2}) = 1 - \alpha.$$

وبإتباع الخطوات الجبرية التي استخدمناها في الفصل الثامن يمكن كتابة :

$$P(B_1 - t_{\alpha/2}\sqrt{\frac{S^2}{SXX}} < \beta_1 < B_1 + t_{\alpha/2}\sqrt{\frac{S^2}{SXX}}) = 1 - \alpha.$$

لعينة عشوائية معطاة من الحجم n نحسب s و SXX وذلك للحصول على $(1-\alpha)100\%$ فترة ثقة كالتالي :

$$b_1 - t_{\alpha/2}\sqrt{\frac{s^2}{SXX}} < \beta_1 < b_1 + t_{\alpha/2}\sqrt{\frac{s^2}{SXX}}.$$

مثال (١٠-٤) أوجد 95% فترة ثقة للمعلمة β_1 في معادلة الانحدار $\mu_{y|x} = \beta_0 + \beta_1 x$ بالاعتماد على البيانات في جدول (١٠-٥) .

الحل . للمثال (١٠-٣) فإن :

$$SXX = 46.91667 \quad , \quad b_1 = 44.41385 \quad , \quad s^2 = 33.6888$$

باستخدام جدول توزيع t في ملحق (٤) فإن $t_{.025}=2.228$ بدرجات حرية $\nu = 10$. وعلى ذلك فإن 95% فترة ثقة للمعلمة β_1 تحسب كالآتي :

$$44.41385 - 2.228\sqrt{\frac{33.6888}{46.91667}} < \beta_1 < 44.41385 + 2.228\sqrt{\frac{33.6888}{46.91667}}$$

أي أن :

$$44.41385 - 2.228(0.847382) < \beta_1 < 44.41385 + 2.228(0.847382)$$

والتي تختصر إلى :

$$42.5 < \beta_1 < 46.3.$$

لاختبار فرض العدم $H_0 : \beta_1 = \beta_1^*$ ضد فرض بديل مناسب يمكننا استخدام توزيع t بدرجات حرية n-2 للحصول على منطقة الرفض. قرارنا سوف يعتمد على القيمة :

$$t = \frac{b_1 - \beta_1^*}{\sqrt{s^2 / SXX}}.$$

مثال (١٠-٥) باستخدام القيمة المقدرة $b_0=44.41385$ في مثال (١٠-٣) ، أختبر فرض العدم أن $\beta_1 = 0$ عند مستوى معنوية $\alpha = 0.05$.

الحل .

$$H_0 : \beta_1 = 0,$$
$$H_1 : \beta_1 \neq 0.$$
$$\alpha = 0.05.$$

$t_{.025}=2.228$ ومنطقة الرفض $T > 2.228$ أو $T < -2.228$

$$t = \frac{b_1 - 0.0}{\sqrt{s^2 / SXX}}$$

$$= \frac{44.41385}{\sqrt{\frac{33.6888}{46.91667}}} = \frac{44.41385}{0.847382} = 52.41.$$

وبما أن t تقع في منطقة الرفض نرفض H_0 .

أيضا المتغير العشوائي B_0 له توزيع طبيعي بمتوسط :

$$\mu_{B_0} = \beta_0$$

وتباين :

$$\sigma_{B_0}^2 = \sigma^2 (\frac{1}{n} + \frac{\bar{x}^2}{SXX})$$

وتبعا لنظرية (٨-١) فإن :

$$Z = \frac{B_0 - \beta_0}{\sqrt{\sigma^2 (\frac{1}{n} + \frac{\bar{x}^2}{SXX})}}$$

متغير عشوائي يتبع التوزيع الطبيعي القياسي وباستبدال σ بالإحصاء S في صيغة Z نحصل على المتغير العشوائي :

$$T = \frac{B_0 - \beta_0}{\sqrt{S^2(\frac{1}{n} + \frac{\overline{x}^2}{SXX})}}$$

والذي يتبع توزيع t بدرجات حرية n −2 .

سوف يستخدم المتغير T للحصول على $100\%(1-\alpha)$ فترة ثقة للمعلمة β_0 كالتالي :

$$b_0 - t_{\alpha/2}\sqrt{s^2(\frac{1}{n} + \frac{\overline{x}^2}{SXX})} < \beta_0 < b_0 + t_{\alpha/2}\sqrt{s^2(\frac{1}{n} + \frac{\overline{x}^2}{SXX})}.$$

مثال (١٠−٦) أوجد 95% فترة ثقة للمعلمة β_0 في خط الانحـــدار $\mu_{y|x} = \beta_0 + \beta_1 x$ بالاعتماد على البيانات في جدول (١٠−٣) .

الحل . من المثال (١٠−٣) فإن :

$$s^2 = 33.6888 \ , \ SXX = 46.91667 \ , \ \overline{x} = 4.58333 \ , \ b_0 = 14.187$$

وعلى ذلك 95% فترة ثقة للمعلمة β_0 تعطى على الشكل :

$$14.187 - 2.228\sqrt{33.6888\left[\frac{1}{12} + \frac{4.58333^2}{46.91667}\right]} < \beta_0$$

$$14.187 + 2.228\sqrt{33.6888\left[\frac{1}{12} + \frac{4.58333^2}{46.91667}\right]}.$$

أي أن :

$$14.187 - 2.228(4.2298) < \beta_0 < 14.187 + 2.228(4.2298)$$

والتي تختزل إلى :

$$4.8 \ < \ \beta_0 \ < \ 23.6.$$

لاختبار فرض العدم $H_0 : \beta_0 = \beta_0^*$ ضد أي فرض مناسب فإننا مرة أخرى ســـوف نستخدم توزيع t بدرجات حرية n-2 للحصول على منطقة الرفض وبالتالي فـــإن قرارنـا سوف يعتمد على القيمة :

$$t = \frac{b_0 - \beta_0^*}{\sqrt{s^2(\frac{1}{n} + \frac{\overline{x}^2}{SXX})}}$$

الطريقة المتبعة موضحة في المثال التالي :

مثال (١٠-٧) باستخدام القيمة المقدرة $b_0=14.187$ في مثال (١٠-٣) ، أختبر فرض العدم $H_0 : \beta_0 = 0$ ضد الفرض البديل $H_0 : B_0 \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

الحل .

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

مستوى معنوية $\alpha = 0.05$

$$t = \frac{b_0 - 0.0}{\sqrt{s^2(\frac{1}{n} + \frac{\bar{x}^2}{SXX})}}$$

$$= \frac{14.187}{\sqrt{33.6888\left[\frac{1}{12} + \frac{4.58333^2}{46.91667}\right]}} = 3.354.$$

$t_{.025}=2.228$ ومنطقة الرفض $T > 2.228$ أو $T < -2.228$ وحيث أن t وقعت في منطقة الرفض نرفض H_0 .

(١٠-٢-٨) التنبــؤ Prediction

يمكن استخدام المعادلة $\hat{y}_x = b_0 + b_1x$ للتنبأ بقيمة $\mu_{Y|x'}$ ، حيث x' ليس من الضرورة أن تكون واحدة من $x_1, x_2, ..., x_n$ في العينة العشوائية من الحجم n للمشاهدات $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. أيضا يمكن استخدام المعادلة $\hat{y}_x = b_0 + b_1x$ للتنبأ بقيمة واحدة $y_{x'}$ للمتغير $Y | x'$. سوف نتوقع أن خطأ التنبأ سوف يكـون أعلـى في حالة قيمة واحدة متنبأ بها عنه في حالة التنبأ بالمتوسط وهذا سوف يؤثر على طول فـترة الثقة للمعالم المراد تقديرها.

عند الرغبة في الحصول على فترة ثقة للمعلمة $\mu_{Y|x'}$ يكون من الضروري إيجـاد التوزيـع العيني للفروق بين القيم $\hat{y}_{x'}$ والتي نحصل عليها من خط الانحدار المقدر بتكرار المعاينة والقيمـة الحقيقية المقابلة $\mu_{Y|x'}$ من خط الانحدار الحقيقي . باستخدام نظريات الإحصاء يمكن إثبات أن التوزيع العيني للإحصاء $\hat{Y}_{x'} - \mu_{Y|x'}$ يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{Y}_{x'} - \mu_{Y|x'}} = E\left[\hat{Y}_{x'} - \mu_{Y|x'}\right] = 0$$

وتباين :

$$\sigma^2_{\hat{Y}_{x'}-\mu_{Y|x'}} = \sigma^2\left[\frac{1}{n}+\frac{(x'-\bar{x})^2}{SXX}\right].$$

في التطبيق يستبدل σ^2 بالقيمة s^2 والتي تمثل قيمة للإحصاء S^2 وعلى ذلك ، فإن الإحصاء :

$$T = \frac{\hat{Y}_{x'} - \mu_{Y|x'}}{\sqrt{S^2\left[\frac{1}{n}+\frac{(x'-\bar{x})^2}{SXX}\right]}}$$

له توزيع t بدرجات حرية n-2. يمكن الحصول على $(1-\alpha)100\%$ فترة ثقة للمعلمة $\mu_{Y|x'}$ من الصيغة التالية :

$$\hat{y}_{x'} - t_{\alpha/2}\sqrt{s^2\left[(\frac{1}{n}+\frac{(x'-\bar{x})^2}{SXX})\right]} < \mu_{Y|x'} < \hat{y}_{x'} + t_{\alpha/2}\sqrt{s^2\left[(\frac{1}{n}+\frac{(x'-\bar{x})^2}{SXX})\right]}.$$

مثال (١٠-٨) باستخدام البيانات في جدول (١٠-٥) أوجد 95% فترة ثقة للمعلمة $\mu_{Y|4}$.

الحل . من معادلة الانحدار المقدرة فإن :

$$\hat{y}_4 = 14.187 + (44.41385)(4)$$
$$= 191.84.$$

عرفنا مما سبق أن :

$$SXX = 46.91667,\ \bar{x} = 4.58333,\ s^2 = 33.6888,$$

$t_{.025} = 2.228$ بدرجات حرية 10 . وعلى ذلك 95% فترة ثقة للمعلمة $\mu_{Y|4}$ هي :

$$191.84 - 2.228\sqrt{33.6888\left[\frac{1}{12}+\frac{(4-4.58333)^2}{46.91667}\right]} < \mu_{Y|4} <$$
$$191.84 + 2.228\sqrt{33.6888\left[\frac{1}{12}+\frac{(4-4.58333)^2}{46.91667}\right]}\ .$$

أي أن :

$$191.84-(2.228)(1.7469) < \mu_{Y|4} < 191.84+(2.228)(1.7469)$$

والتي تختصر إلى :

$$187.94791 < \mu_{Y|4} < 195.73209.$$

للحصول على $100\%(1-\alpha)$ فترة ثقة لأي قيمة مفردة $y_{x'}$ للمتغير $Y \mid x'$ ، يكون من الضروري تقدير التباين للفروق بين القيم $\hat{y}_{x'}$ المحسوبة من خط الانحدار المقدر بتكرار المعاينة والقيمة المقابلة الحقيقية $y_{x'}$. باستخدام نظريات الإحصاء يمكن إثبات أن التوزيع العيني للإحصاء $\hat{Y}_{x'} - y_{x'}$ يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{Y}_{x'} - y_{x'}} = 0$$

وتباين :

$$\sigma^2_{\hat{Y}_{x'} - y_{x'}} = \sigma^2\left[1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX}\right]$$

في التطبيق يستبدل σ^2 بالقيمة s^2 والتي تمثل قيمة للإحصاء S^2 وعلى ذلك ، فإن الإحصاء :

$$T = \frac{\hat{Y}_{x'} - y_{x'}}{\sqrt{S^2\left[1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX}\right]}}.$$

يتبع توزيع **t** بدرجات حرية **n-2**.

يمكن الحصول على $100\%(1-\alpha)$ فترة لقيمة مفردة $y_{x'}$ من الصيغة التالية :

$$\hat{y}_{x'} - t_{\alpha/2}\sqrt{1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX}} < y_{x'} < \hat{y}_{x'} + t_{\alpha/2}\sqrt{1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX}}.$$

مثال (١٠-٩) للبيانات في جدول (١٠-٥) أوجد (ا) 95% فترة ثقة لـ y_4 (ب) وضح بيانيا 95% فترة ثقة لـ $y_{x'}$.

الحل . (ا) $\bar{x} = 4.58333$, $s^2 = 33.6888$, $n = 12$ وعلى ذلك 95% فترة ثقة لـ y_4 هي :

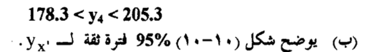
$$191.84 - 2.228\sqrt{33.6888\left[1 + \frac{1}{12} + \frac{(4 - 4.58333)^2}{46.91667}\right]} < y_4 <$$

$$191.84 + 2.228\sqrt{33.6888\left[1 + \frac{1}{12} + \frac{(4 - 4.58333)^2}{46.91667}\right]}$$

أي أن :

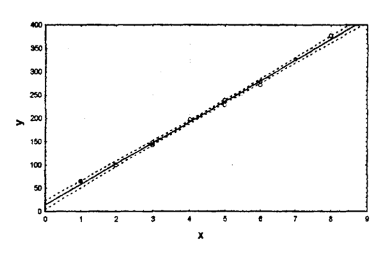
$$\mathbf{191.84 - 2.228(6.061) < y_4 < 191.84 + 2.228\ (6.06)}$$

والتي تختصر إلى :



$$178.3 < y_4 < 205.3$$

(ب) يوضح شكل (١٠-١٠) 95% فترة ثقة لـ $y_{x'}$.



شكل (١٠-١٠)

(١٠-٢-٩) اختبار خطية الانحدار Test for Linearity of Regression

لمشكلة معطاة فإننا نفترض إما أن الانحدار خطي ونتبع خطوات التقدير التي تناولناها في البند (١٠-٢-٣) أو نستنتج أن الانحدار غير خطي وفي هذه الحالة سوف نتبع الخطوات التي سوف نتناولها في البند (١٠-٣). في الجزء التالي سوف نتناول اختبار الخطية في حالتين .

الحالة الأولى عندما تكون σ^2 معلومة (التباين معلوم)

لاختبار فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_1 : النموذج غير خطي فإن:

$$\chi^2 = \frac{SSE}{\sigma^2}$$

قيمة لمتغير عشوائي X^2 له توزيع χ^2 بدرجات حرية $v = n - 2$ وذلك بافتراض صحة فرض العدم. لمستوى معنوية α فإن منطقة الرفض $X^2 > \chi^2_\alpha$ حيث χ^2_α تستخرج من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $v = n - 2$. إذا وقعت χ^2 في منطقة الرفض نرفض H_0.

مثال (١٠-١٠) لأزواج القياسات في جدول (١٠-٧) أختبر فرض العدم H_0 : النمـوذج خطي ضد الفرض البديل H_1 : النموذج غير خطي وذلك عند مسـتوى معنويـة $\alpha = 0.01$ وتحت فرض أن $\sigma^2 = 1$.

جدول (١٠-٧)

x	0.345	0.287	0.251	0.225	0.207	0.186	0.161	0.132	0.084	0.060
y	367	311	295	268	253	239	220	213	193	192

الحل . لاختبار فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_1 : النموذج غير خطي نحسب جدول تحليل التباين والمعطى في جدول (١٠-٨) .ونحسب القيمة :

$$\chi^2 = \frac{SSE}{\sigma^2} = \frac{21.95}{1} = 21.95.$$

بمستوي معنوية $\alpha = 0.01$ فإن $\chi^2_{0.01} = 20.09$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $\nu = n - 2 = 10 - 2 = 8$. منطقة الرفض $X^2 > 20.09$. وبما أن χ^2 تقع في منطقة الرفض نرفض H_0 ، أي أن العلاقة بين المتغيرين لا تتبع النموذج الخطي ويجب أن نبحث عن نموذج آخر.

جدول (١٠-٨)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات
الانحدار	1	342.0	342.0
الخطأ	8	21.95	2.744
الكلي	9	363.95	

الحالة الثانية عندما تكون σ^2 مجهولة :

لاختبار فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_1: النموذج غير خطي نختار عينة عشوائية من الحجم n من المشاهدات بحيث أن لكل قيمة من x يوجد أكثر من قيمة لـ y . أي أن العينة التي حجمها n تحتوي على قيم مختلفة من x عددها k بحيث تحتوي العينة على n_1 من القيم المشاهدة $y_{11}, y_{12}, \dots, y_{1n_1}$ المقابلة للقيمة x_1 و n_2 من القيم المشاهدة $y_{21}, y_{22}, \dots, y_{2n_2}$ المقابلة للقيمة x_2 و... و n_k من القيم المشاهدة $y_{k1}, y_{k2}, \dots, y_{kn_k}$ المقابلة للقيمة x_k حيث $n = \sum_{i=1}^{k} n_i$.

لإجراء الاختبار نتبع الخطوات التالية :

(أ) نحسب مجموع مربعات الخطأ الخالص من x_1 pure error of sum squares من الصيغة التالية :

$$\sum_{u=1}^{n_1} (y_{1u} - \bar{y}_1)^2$$

، حيث $\bar{y}_1 = \sum_{u=1}^{n_1} y_{1u} / n_1$ ، بدرجات حرية n_1-1 . بنفس الشكل يمكن حساب مجموع مربعات الخطأ الخالص من x_2 و ... و x_k . مجموع مربعات الخطأ الخالص الكلي يحسب من الصيغة التالية :

$$SSP = \sum_{i=1}^{k} \sum_{u=1}^{n_i} (y_{iu} - \bar{y}_i)^2$$

بدرجات حرية :

$$n_e = \sum_{i=1}^{k} (n_i - 1) = \sum_{i=1}^{k} n_i - k$$

(ب) نحسب متوسط مجموع مربعات الخطأ الخالص من الصيغة التالية :

$$s_e^2 = \frac{SSP}{n_e}$$

والذي يعتبر تقدير للمعلمة σ^2

(ج) نحسب مجموع مربعات قصور جودة التوفيق **sum squares lake of fit** كالتالي :

$$SSL = SSE - SSP$$

بدرجات حرية $n_L = (n-2) - n_e$.

من الحسابات السابقة فإن جدول تحليل التباين يكون كما هو موضح في جدول (١٠-٩)

جدول (١٠-٩)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات
الانحدار الخطــأ	1 n-2	SSR SSE	SSR/1 SSE/n-2
الكلي	n-1		
قصور جودة التوفيق الخطأ الخالص	n_L n_e	SSL SSP	$MSL=SSL/n_L$ $s_e^2 = SSP/n_e$

نحسب قيمة f' كالتالي :

$$f' = \frac{MSL}{s_e^2}.$$

بافتراض صحة فرض العدم فإن f' تمثل قيمة لمتغير عشوائي **F** له توزيع **F** بدرجات حرية n_L , n_e . عند مستوي معنوية α نحسب منطقة الرفض كالتالي : $F > f_\alpha(n_L, n_e)$ حيــث:

$f_\alpha(n_L, n_e)$ تستخرج من جدول توزيع F في ملحق (٦) أو ملحق (٧) بدرجات حرية $\nu_1 = n_L, \nu_2 = n_e$. إذا وقعت f' في منطقة الرفض نرفض H_0 .

مثال (١٠-١١) لازواج القياسات في جدول (١٠-١٠) اختبر فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_0 : النموذج غير خطي عند مستوى معنوية $\alpha = 0.05$.

جدول (١٠-١٠)

المشاهدات	x	y	المشاهدات	x	y	المشاهدات	x	y
1	1.3	2.3	9	3.7	1.7	17	5.3	3.5
2	1.3	1.8	10	4.0	2.8	18	5.3	2.8
3	2.0	2.8	11	4.0	2.8	19	5.3	2.1
4	2.0	1.5	12	4.0	2.2	20	5.7	3.4
5	2.7	2.2	13	4.7	5.4	21	6.0	3.2
6	3.3	3.8	14	4.7	3.2	22	6.0	3.0
7	3.3	1.8	15	4.7	1.9	23	6.3	3.0
8	3.7	3.7	16	5.0	1.8	24	6.7	5.9

الحل .

H_0 : النموذج الخطى :

H_1 : النموذج غير خطى :

$\alpha = 0.05$.

سوف نوجد مجموع المربعات الخالص ثم مجموع مربعات قصور التوفيق كالتالي :

مجموع مربعات الخطأ الخالص عند $x = 1.3$ هو :

$$(2.3)^2 + (1.8)^2 - \{ (2.3 + 1.8)^2 / 2 \} = 0.125$$

بدرجة حرية واحدة ($n_1 = 2\text{-}1 = 1$) .

مجموع مربعات الخطأ الخالص عند $x = 2.0$ هو :

$$(2.8)^2 + (1.5)^2 - \{ (2.8 + 1.5)^2 / 2 \} = 0.845$$

بدرجة حرية واحدة ($n_2 = 2\text{-}1 = 1$) . بنفس الطريقة يتم حساب مجموع مربعات الخطأ الخالص للقيم الباقية من x كما في جدول (١٠ - ١١) .

جدول (١٠-١١)

مستوى x	$\Sigma(y_{iu} - \bar{y}_i)^2$	درجات حرية
1.3	0.125	1
2.0	0.845	1
3.3	2.000	1
3.7	2.000	1

←

4.0	0.240	2
4.7	6.260	2
5.3	0.980	2
6.0	0.020	1
المجموع	12.470	11

جدول تحليل التباين معطى في جدول (١٠-١٢).

جدول (١٠-١٢)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات	f المحسوبة
الانحدار	1	6.326	6.326	$f=\frac{6.326}{0.963}$ $=6.569$ عند مستوى معنوية $\alpha=0.05$
الخطأ	22	21.192	$0.963=s^2$	
قصور التوفيق	11	8.722	0.793=MSL	$f'=\frac{0.793}{1.134}$ $=0.699$
الخطأ الخالص	11	12.470	$1.134=s_e^2$	

من جدول (١٠-١٢) فإن القيمة $f'=0.699$ غير معنوية لأنها أقل من الواحد.

(١٠-٣) بعض نماذج الانحدار الغير خطية

Some Nonlinear Regression Models

يوجد العديد من النماذج الرياضية الغير خطية المستخدمة في تمثيل العلاقـــات الاقتصاديـــة والاجتماعية ... الخ. سوف تقتصر دراستنا في هذا الجزء على النموذج الأسى ونموذج القــوى.

(١٠-٣-١) النموذج الأسي The Exponential Model

معادلة النموذج الأسي تكون على الصورة التالية :

$$\mu_{Y|x}=\gamma\,\delta^{x}$$

حيث γ,δ ثابتان والمطلوب تقديرهما من البيانات بالتقديرين c , d على التوالي . يمكن تقدير $\mu_{Y|x}$ بالقيمة $\hat{y}_x$ من منحنى الانحدار المقدر التالي :

$$\hat{y}_x=c\ d^{x}.$$

بأخذ لوغاريتمات الطرفين (للأساس e) في المعادلة السابقة فإن منحنى الانحدار المقــدر يمكــن كتابته على الشكل :

$$\ln\hat{y}_x=\ln c+(\ln d)x,$$

وكل زوج من المشاهدات في العينة يحقق العلاقة :

$$\ln y_i = \ln c + (\ln d)\, x_i + e_i = b_0 + b_1 x_i + e_i ,$$

حيث أن : $b_1 = (\ln d), b_0 = \ln c$. وعلى ذلك يمكن إيجاد b_0, b_1 بالصيغ المستخدمة لنموذج الانحدار الخطى ، التي سبق أن تناولناها ، باستخدام النقاط **$(x_i , \ln y_i)$** ثم إيجاد **c , d** بأخذ القيم المقابلة للوغاريتمات لــ b_0, b_1 على التوالي، أي أن :

$$d = \exp(b_1), c = \exp(b_0) .$$

طريقة المربعات الصغرى لتوفيق منحنى أسى لفئة من المشاهدات موضحه في المثال التالي .

مثال (١٠-١٢) لازواج القياسات في جدول (١٠-١٣) أوجد معادلة الانحدار المقدرة تحت فرض النموذج الأسى .

جدول (١٠ - ١٣)

x	1	2	3	4	5	6	7
y	304	341	393	457	548	670	882

الحل . بوضع $y'_i = \ln y_i$ فإن : $\Sigma y'_i = 43.243148$

و $\quad n = 7 \quad , \quad \Sigma x_i = 28 \quad , \quad \Sigma x_i^2 = 140$

$$\Sigma x_i y'_i = 177.85134 \quad , \quad \bar{x} = 4 , \frac{\Sigma y'_i}{n} = 6.1775926$$

$$\mathbf{b_1} = \frac{177.85134 - \frac{(28)(43.243148)}{7}}{140 - \frac{(28)^2}{7}}$$

$$= 0.174241 .$$

$$b_0 = 6.1775926 - (0.174241)(4)$$

$$= 5.4806286 .$$

معادلة الانحدار المقدرة هي :

$$\hat{y}_x = 5.4806286 + 0.174241x.$$

وعلى ذلك :

$$\ln d = b_1 = 0.174241 \ , \ \ln c = b_0 = 5.4806286,$$

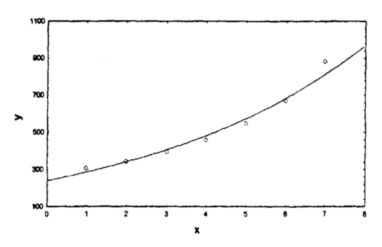
$$d = \exp(b_1) = 1.1903424 , c = \exp(b_0) = 239.99752.$$

وبالتالي فإن منحنى الانحدار المقدر بالمربعات الصغرى هو :

$$\hat{y}_x = c \ d^x$$

$$= (239.99752)(1.1903424)^x$$

والتمثيل البياني لها موضح في شكل (١٠- ١١) .



شكل (١٠- ١١)

(١٠-٣-٢) نموذج القوى **Power Model**

معادلة نموذج القوى تكون على الصورة التالية :

$$\mu_{Y|x} = a_0 \; x^{\,a'_0}$$

حيث a_0 , a'_0 ثابتان والمطلوب تقديرهما من البيانات بالتقديرين c_o , d_o على التوالي. يمكن تقدير $\mu_{Y|x}$ بالقيمة $\hat{y}_x$ من منحنى الانحدار المقدر التالي :

$$\hat{y}_x = c_o \; x^{d_o} .$$

بأخذ لوغاريتمات الطرفين (للأساس e) فإن منحنى الانحدار يمكن كتابته على الشكل :

$$\ln \hat{y}_x = \ln \; c_o + d_o(\ln \; x)$$

كل زوج من المشاهدات في العينة يحقق العلاقة :

$$\ln y_i = \ln \; c_o + d_o(\ln \; x_i) + e_i$$
$$= b_o + b_1(\ln \; x_i) + e_i$$

حيث $b_1 = d_o, b_o = \ln \; c_o$. وعلى ذلك يمكن إيجاد b_0 , b_1 بالصيغ المستخدمة لنموذج الانحدار الخطى، التي سبق أن تناولناها ، باستخدام النقاط $(\ln \; x_i , \ln \; y_i)$ ثم إيجاد c_o , d_o حيث $b_1 = d_o , \ln c_o = b_o$.

مثال (١٠-١٣) لأزواج القياسات في جدول (١٠-١٤) أوجد معادلة الانحدار المقدرة تحت فرض نموذج القوى .

جدول (١٠ – ١٤)

x	600	600	600	600	500	500	500	500	400	400	400	400
y	2.35	2.65	3.0	3.6	6.4	7.8	9.8	16.5	21.5	24.5	26.0	33.0

الحل .

$$n=12\ ,\ \Sigma \ln x_i = 74.412\ ,\ \Sigma \ln y_i = 26.22601\ ,$$

$$\Sigma \ln x_i^2 = 461.75874\ ,\ \Sigma(\ln x_i)(\ln y_i) = 160.84601\ ,$$

$$\Sigma \ln y_i^2 = 67.74609.$$

الحل.

$$b_1 = \frac{160.84601 - \dfrac{(74.412)(26.22601)}{12}}{461.75874 - \dfrac{(74.412)^2}{12}}$$

$$= -5.3996,$$

$$b_0 = \frac{26.22061 - (-5.3996)(74.412)}{12}$$

$$= 35.6684.$$

معادلة الانحدار المقدرة هي :

$$\hat{y}_x = 35.6684 - 5.3996x.$$

وعلى ذلك :

$$\ln c_o = b_0 = 35.6684$$

أي أن :

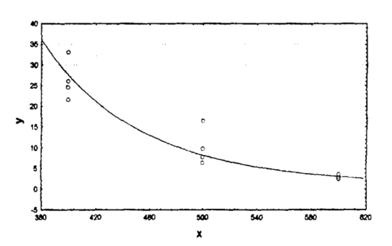
$$c_0 = \exp(b_0) = 3.094491530.10^{15}$$

$$d_0 = b_1 = -5.3996$$

والمعادلة الأساسية المقدرة هي :

$$\hat{y}_x = c_0\ x^{d_0} = (3.094491530 \cdot 10^{15}) \cdot x^{-5.3996}.$$

والتمثيل البياني لها موضح في شكل (١٠– ١٢).



شكل (١٠- ١٢)

(١٠-٤) معامل الارتباط الخطي البسيط

The Simple Linear Correlation Coefficient

في مشكلة الانحدار كان اهتمامنا بالتنبأ بمتغير وذلك من المعلومات عن المتغيرات المستقلة ، بينما في مشكلة الارتباط فإن اهتمامنا سوف يكون في قياس العلاقة بين متغيرين أو أكثر. مرة أخرى فإن قيم المتغيرات المستقلة كانت ثابتة في مشكلة الانحدار .الآن سوف يختلف الوضع. سوف نعرف معامل الارتباط الخطي بأنه مقياس للعلاقة بين متغيرين عشوائيين X,Y. وسوف نرمز له بالرمز r . سوف نفترض أن المتغيران X,Y لهما توزيع احتمالي ثنائي. لحساب معامل الارتباط الخطي نختار عينة عشوائية من أزواج المشاهدات (x , y) . إذا كانت نقاط شكل الانتشار تتركز فوق وحول خط انحدار له ميل موجب ، فهذا يدل على ارتباط موجب قوى بين المتغيرين (ارتباط طردي) كما هو موضح في شكل (١٠-١٣) (a) . ومن ناحية أخرى ، إذا كانت نقاط شكل الانتشار تتركز فوق وحول خط انحدار له ميل سالب فهذا يدل على ارتباط قوى سالب بين المتغيرين (ارتباط عكسي) كما هو موضح في شكل (١٠-١٣) (b) كلما زاد انتشار نقاط شكل الانتشار حول وفوق خط الانحدار فإن الارتباط يقل عدديا بين المتغيرين ، إذا كانت نقاط شكل الانتشار تنتشر بطريقة عشوائية كما في شكل (١٠-١٣) (c) فهذا يعني أن r=0 ونستنتج عدم وجود علاقة بين X,Y. ولما كان معامل الارتباط بين متغيرين يعتبر مقياس للعلاقة الخطيه بينهما وعلى ذلك فإن r=0 تعني قصور في الخطية وليست قصور في الارتباط . على سبيل المثال قد تكون هناك علاقة ولكنها علاقة غير خطيه . فعلى سبيل المثال إذا

وجدت علاقة قوية من الدرجة الثانية بين المتغيرين X,Y كما هو موضح في شكل (١٠-١٣) (d) فهذا يعني أن r=0 . يعتبر معامل الارتباط الخطي (معامل بيرسون للارتباط) أو اختصارا معامل الارتباط أكثر مقاييس الارتباط الخطي انتشارا.

y x (a) y x (b)

y x (c) y x (d)

شكل (١٠-١٣)

يتم حساب معامل الارتباط من المعادلة التالية :

$$r = \frac{\Sigma x_i y_i - \frac{\Sigma x_i y_i}{n}}{\sqrt{\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right]\left[\Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}\right]}}$$

$$= \frac{SXY}{\sqrt{SXX.SYY}}$$

من الجزء (١٠-٢-٦) يمكن حساب معامل الارتباط كالتالي :

$$r = \pm\sqrt{r^2}$$

حيث r^2 هو معامل التحديد البسيط والإشارة تخص التقدير b_1 . وبما أن $0 \le r^2 \le 1$ فــإن r يمكن أن تأخذ الإشارة الموجبة أو السالبة ، أي أن :

$$-1 \le r \le 1$$

عادة يفضل حساب معامل الارتباط من معادلته وليس من r^2 لصعوبة الحساب.

مثال (١٠-١٤) لدراسة العلاقة بين تركيز الأوزون Ozone (X) (مقاس PPM) وتركيز الكربون (Y) (مقاس $\mu g / m^3$) تم الحصول على البيانات المعطاة في جدول (١٠-١٥).

جدول (١٠-١٥)

x	0.066	0.088	0.120	0.050	0.162	0.186	0.057	0.100
y	4.6	11.6	9.5	6.3	13.8	15.4	2.5	11.8
x	0.112	0.055	0.154	0.074	0.111	0.140	0.071	0.110
y	8.0	7.0	20.6	16.6	9.2	17.9	2.8	13.0

أوجد معامل الارتباط البسيط .

الحل .

$$n = 16 \ , \ \Sigma x_i = 1.656 \ , \ \Sigma y_i = 170.6 \ ,$$

$$\Sigma x_i^2 = 0.196912 \ , \ \Sigma x_i \ y_i = 20.0397 \ ,$$

$$\Sigma y_i^2 = 2253.56.$$

$$SXY = \Sigma x_i \ y_i - \frac{\Sigma x_i \ \Sigma y_i}{n}$$

$$= 20.0397 - \frac{(1.656)(170.6)}{16}$$

$$= 2.3826,$$

$$SXX = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

$$= 0.196912 - \frac{(1.656)^2}{16} = 0.025516,$$

$$SYY = \Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n} = 2253.56 - \frac{(170.6)^2}{16}$$

$$= 434.5375.$$

وعلى ذلك :

$$r = \frac{SXY}{\sqrt{SXX.SYY}} = \frac{2.3826}{\sqrt{(0.025516)(434.5375)}} = 0.716.$$

في المناقشة السابقة لم نضع فروض قوية على توزيع المجتمع الذي اختبرت منه العينة وذلك للحصول على تقدير بنقطة للمعلمة ρ والتي ترمز إلى معامل ارتباط المجتمع. للحصول على

$100\%(1-\alpha)$ فترة ثقة للمعلمة ρ أو اختبارات فروض تخص ρ فإننا نفترض أن العينة مأخوذة من مجتمع يتبع التوزيع الطبيعي الثنائي **the bivarate normal distribution** أي أن X , Y متغيرين عشوائيين حيث دالة التوزيع الهامشية لكل من Y , X تتبع التوزيع الطبيعي . يمكن اختبار الاعتدال لقيم X وقيم Y على حدة بطريقة رياضية سوف نتناولها في البند (١٢–٥) من الفصل الثاني عشر .

<u>اختبارات فروض وفترات ثقة تخص ρ</u>

<u>Tests Hypotheses and Confidence Intervals Concerning</u> ρ

لاختبار فرض العدم $H_0: \rho = 0$ ضد الفرض البديل $H_1 : \rho \neq 0$ أو الفرض البديل $H_1 : \rho > 0$ أو الفرض البديل $H_1 : \rho < 0$ وبافتراض صحة فرض العدم فإن :

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

هي قيمة لمتغير عشوائي T له توزيع t بدرجات حرية $v = n - 2$. وعلى ذلك لمستوى معنوية α وللفرض البديل $H_1 : \rho \neq 0$ (اختبار ذي جانبين) فإن منطقة الرفض سوف تكون $T < -t_{\alpha/2}$ or $T > t_{\alpha/2}$ حيث $t_{\alpha/2}$ هي قيمة t المستخرجة من جدول توزيع t بدرجات حرية $v = n - 2$. للبديل $H_1 : \rho > 0$ فإن منطقة الرفض $T > t_{\alpha}$ وللبديل $H_1 : \rho < 0$ فإن منطقة الرفض $T < -t_{\alpha}$.

مثال (١٠–١٥) بفرض أن البيانات في مثال (١٠–١٤) مأخوذة من مجتمع يتبع التوزيع الثنائي الطبيعي . المطلوب اختبار فرض العدم $H_0: \rho = 0$ ضد الفرض البديل $H_1 : \rho > 0$ وذلك عند مستوى معنوية $\alpha=0.01$.

الحل .

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.716\sqrt{14}}{\sqrt{1-(0.716)^2}} = 3.84.$$

$t_{0.01}= 2.624$ والمستخرجة من جدول توزيع t في ملحق (٤) بدرجات حرية $v = n - 2 = 16 - 2 = 14$. منطقة الرفض $T > 2.624$. وبما أن t تقع في منطقة الرفض ، نرفض H_0 .

وبما أن ρ يقيس قوة الارتباط الخطى بين متغيرين X,Y في المجتمع فإن فرض العدم $H_0: \rho=0$ يدل على عدم وجود ارتباط بين المتغيرين في المجتمع. في البند (١٠–٢–٧) استخدمنا

القيمة $t = \frac{b_1}{\sqrt{s^2 / SXX}}$ لاختبار فرض العدم $H_0: \beta_1 = 0$ هنا يمكن إثبات أن $t = r\sqrt{n-2}/\sqrt{1-r^2} = b_1/\sqrt{s^2/SXX}$ وهذا يعنى أن الاختبارين متكافئين . وعلى ذلك إذا كان الاهتمام فقط بقياس قوة العلاقة بين متغيرين X , Y وليس الحصول على معادلة الانحدار الخطى فإن اختبار $H_0: \rho = 0$ يكون اسهل من اختبار t لأنه يتطلب كمية قليلة من الحسابات.

الأسلوب المستخدم لاختبار $H_0: \rho = \rho_0$ عندما $\rho_0 \neq 0$ لا يكافئ أي طريقة مستخدمة في تحليل الانحدار. بفرض أن أزواج المشاهدات $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ تمثل عينة عشوائية مأخوذة من مجتمع يتبع التوزيع الثنائي الطبيعي وإذا كانت n كبيرة وبافتراض صحة فرض العدم فإن :

$$v = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right)$$

هي قيمة لمتغير عشوائي V تقريباً يتبع التوزيع الطبيعي بمتوسط $\mu_V = \frac{1}{2} \ln\left(\frac{1+\rho_0}{1-\rho_0}\right)$ وتباين $\sigma_V^2 = \frac{1}{n-3}$ ، حيث التباين لا يعتمد على ρ، وعلى ذلك فإن :

$$z = \frac{v - \frac{1}{2} \ln\left(\frac{1+\rho_0}{1-\rho_0}\right)}{1/\sqrt{n-3}}.$$

هي قيمة لمتغير عشوائي Z تقريبا يتبع التوزيع الطبيعي القياسي. الجدول (١٠-١٦) يعطى الفروض البديلة ومنطقة الرفض لكل فرض بديل عند مستوى معنوية α.

جدول (١٠-١٦)

الفروض البديلة	منطقة الرفض
$H_1: \rho \neq \rho_0$	$Z < -z_{\alpha/2}$ or $Z > z_{\alpha/2}$.
$H_1: \rho > \rho_0$	$Z > z_\alpha$
$H_1: \rho < \rho_0$	$Z < -z_\alpha$

مثال (١٠-١٦) إذا كان لديك البيانات التالية :

$$n = 20 \;,\; \Sigma y_i = 690.30 \;,\; \Sigma y_i^2 = 29040.29 \;,$$
$$\Sigma x_i y_i = 10818.56 \;,\; \Sigma x_i = 285.90 \;\; \Sigma x_i^2 = 4409.55,$$

أختبر الفرض $0.8 > \rho > 0.5$ عند مستوى معنوية $\alpha = 0.05$

الحل . أي أننا نرغب في اختبار :

$$H_0 : \rho = 0.5 \,,$$
$$H_1 : \rho > 0.5$$
$$\alpha = 0.05.$$

وحيث أن $r = .733$ فإن :

$$v = \frac{1}{2}\ln\left(\frac{1+.733}{1-.733}\right) = .935,$$
$$\mu_V = \frac{1}{2}\ln\left(\frac{1+.5}{1-.5}\right) = .549.$$

وعلى ذلك فإن :

$$z = \frac{v - \frac{1}{2}\ln[(1+\rho_0)/(1-\rho_0)]}{1/\sqrt{n-3}}$$
$$= (.935 - .549)\sqrt{17} = 1.59.$$

$z_{0.05} = 1.645$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) . منطقة الرفـــض $Z > 1.645$. وبما أن z تقع في منطقة القبول نقبل H_0 .

يمكن الحصول على $100\%(1-\alpha)$ فترة ثقة للمعلمة ρ من الصيغة التالية :

$$\frac{e^{2c_1} - 1}{e^{2c_1} + 1} \le \rho \le \frac{e^{2c_2} - 1}{e^{2c_2} + 1}$$

حيث أن $c_1 = v - \frac{z_{\alpha/2}}{\sqrt{n-3}}$, $c_2 = v + \frac{z_{\alpha/2}}{\sqrt{n-3}}$

للبيانات في مثال (١٠-١٦) فإن :

$$r = 0.733 \quad , \quad v = 0.935 \quad , \quad n = 20,$$
$$c_1 = .935 - 1.96/\sqrt{17} = .460,$$

$$c_2 = .935 + 1.96/\sqrt{17} = 1.410$$

بالتعويض في الصيغة التالية يمكن الحصول على 95% فترة ثقة للمعلمة ρ كالتالي :

$$\frac{e^{2(.460)}-1}{e^{2(.460)}+1} \le \rho \le \frac{e^{2(1.410)}-1}{e^{2(1.410)}+1}$$

والتي تختزل إلى :

$$0.43 \le \rho \le 0.89$$

(١٠-٥) الانحدار الخطي المتعدد Linear Multiple Regression

في الغالب تكون العلاقات الفعلية سواء الاقتصادية أو الاجتماعية أو السياسية معقدة يمثل فيها متغير واحد تابع وعدد من المتغيرات الأساسية المستقلة ومن الأمثلة العديدة على ذلك في مجال الاقتصاد نجد أن الكمية المستهلكة من سلعة ما تتأثر بسعر السلعة ذاتها علاوة على أسعار السلع البديلة وأيضا بالإضافة إلى ذوق المستهلك. كذلك كمية الإنتاج تتأثر بالعمل ورأسي المال والموارد الوسيطية وغيرها من عناصر العملية الإنتاجية. وفي مجال التأمين يتوقف القسط التأميني على عمر المؤمن ودخله وقيمة الوثيقة وطول فترات التأمين .

(١٠-٥-١) طريقة المربعات الصغرى Least Square Method

الآن سوف نتناول مشكلة التقدير والتنبأ بقيمة متغير تابع بالاعتماد على فئة من المشاهدات المأخوذة من عدة متغيرات مستقلة $X_1, X_2, \ldots, X_p$. كما في حالة الانحدار الخطي البسيط ، القيمة لكل متغير مستقل والمختارة بواسطة الباحث سوف تظل ثابتة . إذا تم اختيار عينة عشوائية من الحجم n من المجتمع فإن بيانات العينة سوف تكون على الشكل :

$$\{x_{1i}, x_{2i}, \ldots, x_{pi}; y_i\}; i = 1,2,\ldots,n$$

مرة أخرى القيمة y_i تمثل قيمة لمتغير عشوائي Y_i . نموذج الانحدار الخطي المتعدد النظري سوف يكون على الشكل :

$$\mu_{Y|x_1,x_2,\ldots,x_p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p,$$

حيث أن $\beta_0, \beta_1, \ldots, \beta_p$ تمثل المعالم المطلوب تقديرها. نموذج الانحدار الخطي المتعدد المقدر سوف يكون على الشكل :

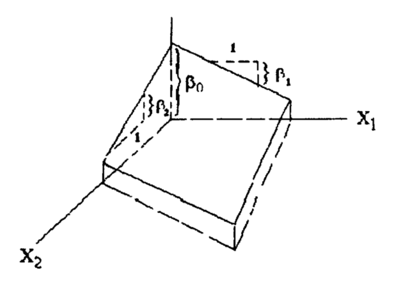
$$\hat{y}_{x_1,x_2,\ldots,x_p} = b_0 + b_1 x_1 + \ldots + b_p x_p,$$

حيث أن $b_0, b_1, \ldots b_p$ التقديرات المطلوب الحصول عليها للمعالم $\beta_0, \beta_1, \ldots, \beta_p$. سوف تقتصر دراستنا في هذا البند على حالة وجود متغيرين مستقلين X_1, X_2 $(p=2)$. النتائج يمكن تعميمها إلى عدة متغيرات مستقلة . في حالة وجود أكثر من متغير مستقل فإن معرفتنا لنظرية المصفوفات سوف يساعدنا في عملية الحساب .

في حالة وجود متغيرين مستقلين فإن نموذج الانحدار الخطي المتعدد سوف يكون على الشكل :

$$\mu_{Y|x_1,x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 ,$$

الشكل البياني للمعادلة السابقة موضح في شكل (١٠-١٤) و الذي يسمى المستوى .



شكل (١٠-١٤)

ونموذج الانحدار الخطي المقدر سوف يكون على الشكل :

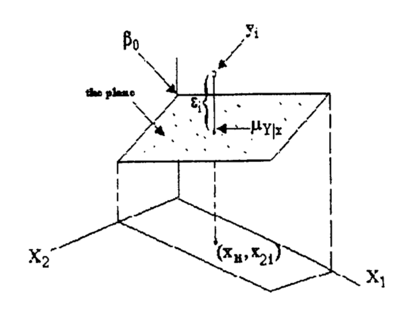
$$\hat{y}_{x_1,x_2} = b_0 + b_1 x_1 + b_2 x_2 ,$$

وكل فئة من المشاهدات تحقق العلاقة :

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + e_i , i = 1,2,\ldots,n,$$

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i , i = 1,2,\ldots,n.$$

الشكل البياني للمعادلة السابقة موضح في شكل (١٠-١٥) .



شكل (١٠-١٥)

تقديرات المربعات الصغرى b_0, b_1, b_2 يمكن الحصول عليها بحل المعادلات الخطية التالية آنيا :

$$b_0 n + b_1 \Sigma x_{1i} + b_2 \Sigma x_{2i} = \Sigma y_i$$

$$b_0 \Sigma x_{1i} + b_1 \Sigma x_{1i}^2 + b_2 \Sigma x_{1i} x_{2i} = \Sigma x_{1i} y_i$$

$$b_0 \Sigma x_{2i} + b_1 \Sigma x_{1i} x_{2i} + b_2 \Sigma x_{2i}^2 = \Sigma x_{2i} y_i$$

مثال (١٠-١٧) يتأثر الاستهلاك السنوي للطعام على كل من الدخل السنوي للأسرة x_1 وحجم الأسرة x_2 . أوجد معادلة الانحدار الخطى المقدرة وذلك بفرض توفر البيانات المعطاة في جدول (١٠-١٧) .

الحل . من البيانات في جدول (١٠-١٧) فإن :

$$n = 10 \ , \ \Sigma x_{1i} = 60 \ , \ \Sigma x_{2i} = 40 \ ,$$

$$\Sigma x_{1i}^2 = 406 \ , \ \Sigma x_{1i} x_{2i} = 269 \ \ \Sigma x_{2i}^2 = 182,$$

$$\Sigma y_i = 180 \ , \ \Sigma x_{1i} y_i = 1159 \ \ \Sigma x_{2i} y_i = 766, \Sigma y_i^2 = 3396.$$

بالتعويض بالقيم السابقة في المعادلات الثلاثة نحصل على :

$$10\, b_0 + 60\, b_1 + 40\, b_2 = 180$$

$$60\, b_0 + 406\, b_1 + 269\, b_2 = 1159$$

$$40\, b_0 + 269\, b_1 + 182\, b_2 = 766.$$

الحل . لهذه الفئة من المعادلات نحصل على :

$$b_0 = 7.918 \ , \ b_1 = 2.363 \ , b_2 = - 1.024$$

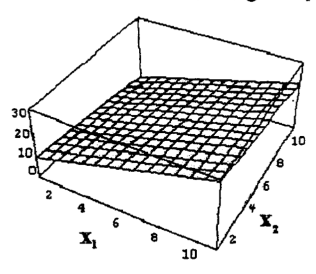
جدول (١٠-١٧)

الأسـرة	الاستهلاك السنوي للطعام (بمئات الدولارات)	الدخل السنوي الصافي (بمئات الدولارات)	حجم الأسرة (عدد الأفراد في الأسرة الواحدة)
1	22	8	6
2	23	10	7
3	18	7	5
4	9	2	2
5	14	4	3
6	20	6	4
7	21	7	4
8	18	6	3
9	16	4	3
10	19	6	3

وعلى ذلك فإن نموذج الانحدار الخطى المقدر يمكن كتابته على الشكل :

$$\hat{y}_{x_1,x_2} = 7.918 + 2.363x_1 - 1.024x_2$$

والتمثيل البياني لها موضح في شكل (١٠-١٦) .



شكل (١٠-١٦)

(١٠-٥-٢) تحليل الانحدار Analysis of Variance

عادة ، الخطوة الأولى بعد الحصول على معادلة الانحدار المتعدد هو اختبار ما إذا كانت هناك علاقة بين المتغير التابع وفئة المتغيرات المستقلة. يعتبر اختبار F في الانحدار المتعدد تعميم لاختبار F في حالة الانحدار الخطي البسيط في حالة متغيرين مستقلين . فرض العدم سوف يكون :

$$H_0 : \beta_1 = 0, \beta_2 = 0$$

ضد الفرض البديل : على الأقل واحد من $\beta_i \neq 0$ و $H_1 : i = 1,2$

يعطى جدول (١٠-١٨) مجاميع المربعات ودرجات الحرية المقابلة لها .

جدول (١٠ - ١٨)

مجموع المربعات	درجات الحرية
$SSTO = \Sigma(y_i - \bar{y})^2$	n-1
$SSE = \Sigma(y_i - \hat{y}_i)^2$	n-3
$SSR = \Sigma(\hat{y}_i - \bar{y})^2$	2

حيث : $\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i}$

جدول تحليل التباين معطى في جدول (١٠-١٩).

جدول (١٠-١٩)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات
الانحدار	2	SSR	$MSR = \dfrac{SSR}{2}$
الخطأ	n-3	SSE	$MSE = \dfrac{SSE}{n-3}$
الكلي	n-1	SSTO	

بافتراض صحة فرض العدم فإن :

$$f = \frac{MSR}{MSE}$$

تمثل قيمة لمتغير عشوائي F يتبع توزيع F بدرجات حرية $v_2 = n - 3$, $v_1 = 2$. لمستوى معنوية α فإن منطقة الرفض $F > f_\alpha(v_1, v_2)$ حيث $f_\alpha(v_1, v_2)$ تستخرج من الجدول في ملحق (٦) بدرجات حرية v_1, v_2 . إذا وقعت f في منطقة الرفض نرفض H_0 .

للبيانات في مثال (١٠-١٧) فرض العدم سوف يكون :

$$H_0 : \beta_1 = 0, \beta_2 = 0$$

ضد الفرض البديل : على الأقل واحد من $0 \neq \beta_i$ و $H_1 : i = 1,2$
يمكن وضع فرض العدم والفرض البديل على الشكل :

H_0 : الانحدار غير معنوي .

H_1 : الانحدار معنوي .

الحل . من البيانات في جدول (١٠-١٧) فإن جدول تحليل التباين معطى في جدول (١٠-٢٠) .

جدول (١٠-٢٠)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات
الانحدار	2	139.56725	69.78363
الخطأ	7	16.43275	2.34754
الكلي	9	156	

$$f=\frac{69.78363}{2.34754}=29.7263.$$

$f_{0.05}(2,7)=4.74$ والمستخرجة من جدول توزيع F في ملحق (٦) بدرجات حرية $v_1 = 2$, $v_2 = 7$. وبما أن f تقع في منطقة الرفض نرفض H_0 .

يعتبر MSE تقدير للتباين σ^2. وعلى ذلك فإن التقدير بنقط للمعلمة σ^2 هو $s^2=2.34754$. وعلى ذلك فإن التقدير بنقطة للانحراف المعياري σ هو :

$$s=\sqrt{2.34754}=1.5321684.$$

(١٠-٥-٣) معامل التحديد المتعدد Coefficient of Multiple Determination

معامل التحديد المتعدد ، يرمز له بالرمز R^2 هو :

$$R^2=\frac{SSR}{SSTO}=1-\frac{SSE}{SSTO}.$$

في حالة وجود متغير مستقل واحد فإن R^2 يصبح معامل التحديد البسيط r^2 . يتراوح قيمة R^2 من الصفر إلى الواحد الصحيح ، أي أن :

$$0 \leq R^2 \leq 1$$

عندما $R^2 = 0$ فهذا يعني أن $b_1 = b_2 = 0$ وعندما $R^2 = 1$ فهذا يعني أن جميع القيم المشاهدة y_i تقع على المستوى المقدر.

(١٠-٥-٤) الارتباط المتعدد والجزئي Multiple and Partial Correlation

الجذر التربيعي الموجب لمعامل التحديد المتعدد R^2 يسمى الارتباط المتعدد ويرمز له بالرمز R ، حيث :

$$R = +\sqrt{R^2}$$

للمثال (١٠-١٧) فإن :

$$R = +\sqrt{0.89466} = 0.94586$$

عموما يقيس معامل التحديد المتعدد R العلاقة بين المتغير التابع Y والمتغيرات المستقلة كلها مجتمعة. بفرض أننا نرغب في إيجاد العلاقة بين المتغير التابع Y وأحد العوامل فقط (بفرض أن العوامل الأخرى ثابتة (أي بحذف تأثير المتغيرات الأخرى) ، وهنا نستخدم معامل الارتباط الجزئي coefficient of partial correlation .

فإذا كنا نرغب في إيجاد معامل الارتباط الجزئي بين المتغير Y والمتغير X_1 مع استبعاد أثر المتغير X_2 فإن معامل الارتباط الجزئي ، يرمز له بالرمز $r_{y1.2}$ هو :

$$r_{y1.2} = \frac{r_{y1} - r_{y2}\ r_{12}}{\sqrt{1 - r_{y2}^2}\sqrt{1 - r_{12}^2}}$$

حيث r_{y1} هو معامل الارتباط البسيط بين المتغير Y والمتغير X_1 وبالمثل يكون y_{y2} و r_{12}.
أيضا معامل الارتباط الجزئي بين المتغير Y والمتغير X_2 مع استبعاد أثر المتغير X_1 ، يرمز له بالرمز $r_{y2.1}$ ، هو :

$$r_{y2.1} = \frac{r_{y2} - r_{y1}\ r_{12}}{\sqrt{1 - r_{y1}^2}\sqrt{1 - r_{12}^2}}$$

معامل الارتباط الجزئي يمكن أن يكون موجباً أو سالباً حيث تقع قيمته في الفترة [1 , 1-] .
لحساب معاملات الارتباط الجزئية من مثال (١٠-١٧) نقوم أولاً بحساب معاملات الارتباط البسيطة التالية :

$$r_{y1} = \frac{\Sigma\ x_{1i}\ y_i - \frac{\Sigma\ x_{1i}\ \Sigma\ y_i}{n}}{\sqrt{\left[\Sigma\ x_{1i}{}^2 - \frac{(\Sigma\ x_{1i})^2}{n}\right]\left[\Sigma\ y_i^2 - \frac{(\Sigma\ y_i)^2}{n}\right]}}$$

$$= \frac{1159 - \frac{(60)(180)}{10}}{\sqrt{\left[406 - \frac{(60)^2}{10}\right]\left[3396 - \frac{(180)^2}{10}\right]}} = 0.9325795,$$

$$r_{y2} = \frac{\Sigma\, x_{2i}\, y_i - \frac{\Sigma\, x_{2i}\, \Sigma\, y_i}{n}}{\sqrt{\left[\Sigma\, x_{2i}{}^2 - \frac{(\Sigma\, x_{2i})^2}{n}\right]\left[\Sigma\, y_i^2 - \frac{(\Sigma\, y_i)^2}{n}\right]}}$$

$$= \frac{766 - \frac{(40)(180)}{10}}{\sqrt{\left[182 - \frac{(40)^2}{10}\right]\left[3396 - \frac{(180)^2}{10}\right]}} = 0.785207,$$

$$r_{12} = \frac{\Sigma\, x_{1i}\, x_{2i} - \frac{\Sigma\, x_{1i}\, \Sigma\, x_{2i}}{n}}{\sqrt{\left[\Sigma\, x_{1i}{}^2 - \frac{(\Sigma\, x_{1i})^2}{n}\right]\left[\Sigma\, x_{2i}^2 - \frac{(\Sigma\, x_{2i})^2}{n}\right]}}$$

$$= \frac{269 - \frac{(60)(40)}{10}}{\sqrt{\left[406 - \frac{(60)^2}{10}\right]\left[182 - \frac{(40)^2}{10}\right]}} = 0.9116072.$$

وعلى ذلك فإن :

$$r_{y1.2} = \frac{r_{y1} - r_{y2}\, r_{12}}{\sqrt{1 - r_{y2}^2}\sqrt{1 - r_{12}^2}}$$

$$= \frac{0.9325795 - (0.785207)(0.9116072)}{\sqrt{(1-(0.7895207)^2}\sqrt{1-(0.9116072)^2}} = 0.859283884.5,$$

$$r_{y2.1} = \frac{r_{y2} - r_{y1}\ r_{12}}{\sqrt{1-r_{y1}^2}\sqrt{1-r_{21}^2}}$$

$$= \frac{0.785207 - (0.9325795)(0.9116072)}{\sqrt{(1-(0.9325795)^2}\sqrt{1-(0.9116072)^2}} = -0.4376576.$$

(١٠-٦) الانحدار من الدرجة الثانية Quadratic Regression

في بعض الأحيان تكون العلاقة بين متغيرين على شكل منحنى من الدرجة الثانية فعلى سبيل المثال بفرض أننا نرغب في تقدير معالم النموذج :

$$\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_2 x^2$$

في الحقيقة نرغب في تقدير معالم النموذج انحدار خطى متعدد على الشكل :

$$\mu_{Y|x} = b_0 + b_1 x_1 + b_2 x_2$$

وذلك بوضع $x_1 = x$, $x_2 = x^2$ في المعادلة السابقة .

مثال (١٠-١٨) لازواج القياسات المعطاة في جدول (١٠ - ٢١) أوجد الانحدار المقدرة لنموذج انحدار من الدرجة الثانية .

جدول (١٠-٢١)

x	1	2	3	4	5	6	7	8	9	10
y	5.0	10.20	15.35	20.50	25.95	32.20	38.50	46.00	53.80	62.00

الحل . من البيانات في جدول (١٠ -٢١) فإن :

$$n = 10 \ , \ \Sigma x_i = 55 \ , \ \Sigma y_i = 309.5 \ ,$$

$$\Sigma x_i y_i = 2218.1 \ , \ \Sigma x_i^2 = 385,$$

$$\Sigma x_i^2 y_i = 17708.2 \ , \ \Sigma x_i^3 = 3025 \ , \ \bar{y} = 30.95,$$

$$\Sigma x_i^4 = 25333 \ , \ \Sigma y_i^2 = 12831.845 \ ,$$

تستخدم القيم السابقة في الحصول على المعادلات التالية :

$$10\, b_0 + 55\, b_1 + 385\, b_2 = 309.5$$
$$55\, b_0 + 385\, b_1 + 3025\, b_2 = 2218.1$$
$$385\, b_0 + 3025\, b_1 + 25333\, b_2 = 17708.2$$

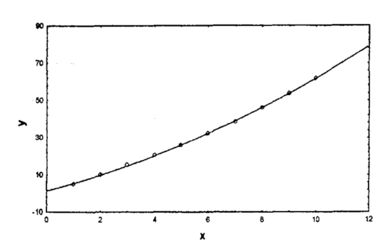
بحل المعادلات السابقة آنيا يمكن إيجاد b_0, b_1, b_2 .

الحل لهذه الفئة من المعادلات هو $b_0 = 1.48083$ و $b_1 = 3.792313$ و $b_2 = 0.223674$

معادلة الانحدار المقدرة هي :

$$\hat{y}_x = 1.48083 + 3.792313x + 0.223674x^2.$$

والتمثيل البياني لها موضح في شكل (١٠- ١٧) .



شكل (١٠- ١٧)

معامل الارتباط من الدرجة الثانية Second – Degree Correlation Coefficient

عندما تكون العلاقة بين متغيرين على شكل معادلة من الدرجة الثانية ، بمعني أن خط الانحدار يكون على شكل منحني (أي غير مستقيم) ، يقال في هذه الحالة أن الارتباط غير مستقيم وفي هذه الحالة لا يصلح قياس الارتباط بمعامل الارتباط الخطي البسيط r وذلك لعدم استقامة الارتباط بل يستخدم المقياس التالي :

$$r = \sqrt{\frac{b_0\Sigma y_i + b_1\Sigma x_i y_i + b_2\Sigma x_i^2 y_i - n\bar{y}^2}{\Sigma y_i^2 - n\bar{y}^2}}$$

$$= \sqrt{\frac{1.4808(309.5) + 3.7923(2218.1) + 0.22367(17708.2) - 10(30.95)^2}{12831.845 - 10(30.95)^2}}$$

$$= \sqrt{\frac{3251.7763}{3252.82}} \quad = \quad \sqrt{0.9997}$$

$= 0.99984.$

أيضا معامل التحديد مازال r^2 يساوى $r^2 = 0.9997$:

تمـــارين :

-١- أجريت تجربة لدراسة تأثير درجة الحرارة X على نتائج إحدى العمليـــات الكيميائيـــة وتم الحصول على البيانات التالية (في شكل شفرة Coded) .

x	-5	-4	-3	-1	-1	0	1	2	3	4	5
y	1	5	4	7	10	8	9	13	14	13	18

أوجد نموذج الانحدار الخطى المقدر .

-٢- تعتبر كمية الرطوبة في منتج ما لها تأثير على كثافة المنتج النهائية. تم مراقبة المنتج وقيـــاس كثافته وتسجيل البيانات التالية (في شكل شفرة Coded) .

رطوبـــة المنتج	x	4.7	5.0	5.2	5.2	5.9	4.7	5.9	5.2	5.9	5.6	5.0
كثافـــة المنتج	y	3	3	4	10	2	9	3	7	6	6	4

(أ) قدر معالم نموذج الانحدار الخطى البسيط .

(ب) أوجد 95% فترة ثقة للمعلمة β_0 .

(ج) أوجد 95% فترة ثقة للمعلمة β_1.

(د) هل تعتقد أن المعادلة الخطية مناسبة لتوفيق هذه البيانات ؟

-٣- إذا كانت تكاليف صيانة سيارات الشحن تزيد مع عمر السيارة . اســـتخدام البيانـــات التالية في (أ) تقدير معالم نموذج الانحدار الخطي البسيط (ب) هل يعتبر النمـــوذج الخطـــي هـــو الأفضل لتوفيق العلاقة في هذا المثال ؟

العمــــر بالسنوات	x	4.5	4.5	4.5	4.0	4.0	5.0	5.0	5.5	5.0	5.0	0.5	6.0	6.0
التكـــاليف خـــلال ٦ شهور	y	619	1049	1033	495	723	681	890	1522	987	1194	163	182	764

-٤- في دراسة عن تأثير درجة الحرارة ، في عملية decolorizing على لون المنتج النـــهائي تم الحصول على البيانات التالية :

درجـــة الحرارة	x	460	450	440	430	420	410	450	440	430

←

اللون	y	0.3	0.3	0.4	0.6	0.5	0.6	0.6	0.5	0.5

(أ) أوجد معادلة نموذج الانحدار الخطى المقدرة.

(ب) أختبر معنوية كل من β_0,β_1 .

(ج) أوجد 95% فترة ثقة لكل من β_0,β_1.

-٥- في أحد أماكن بيع السيارات كانت المبيعات كالتالي :

عمر السيارة	x	3	2	1	1	5	6	1	4
ثمن البيع	y	31	10	59	68	16	15	69	28

(أ) أوجد معادلة الانحدار الخطى المقدرة .

(ب) أوجد 95% فترة ثقة لكل من β_0,β_1.

-٦- إذا كان معروف أن هناك علاقة بين فترة الإصابة بمرض معين عدد البكتريا في العضو المصاب . تم اختيار 10 مصابا بهذا المرض وسجلت أطوال فترات إصابتهم بالمرض عند بدا دخولهم المستشفى فتم الحصول على البيانات التالية :

عدد البكتريا (بالألف)	x	9	10	5	7	10	6	7	4	8	6
فترة الإصابة (باليوم)	y	12	11	8	9	13	10	14	8	11	7

أوجد معادلة الانحدار المقدرة ؟

-٧- فيما يلي أوزان وأطوال مجموعة من الذكور البالغين والمطلوب إيجاد معادلة خط الانحدار المقدرة .

الطول	x	159	180	175	150	170	171	165	176
الوزن	y	68	88	79	65	70	73	63	74

-٨- الجدول التالي يمثل الدخل والأنفاق لعينة من الأسر .

الدخل بمئات الجنيهات	x	42	65	41	43	37	26	38	39
الإنفاق بمئات الجنيهات	y	25	37	25	21	18	24	19	26

(أ) أوجد معادلة الانحدار المقدرة.

(ت) أختبر معنوية β_0 عند مستوى معنوية 0.05.

(ج) أختبر معنوية β_1 عند مستوى معنوية 0.05.

(د) أوجد 95% فترة ثقة لكل من β_0, β_1.

-٩- الجدول التالي يوضح السن وضغط الدم لعشرة من الإناث .

السن x	41	35	62	52	41	58	48	67	66	68
ضغط الدم y	124	115	138	149	145	144	145	152	150	150

(أ) أوجد معادلة الانحدار المقدرة.

(ب) أوجد 99% فترة ثقة لكل من $\beta_0, \beta_1, y_{69}, \mu_{Y|170}$

-١٠- أوضحت الدراسة أن عدد العلب الصفيح التي تتعرض للتلف في عربة الشحن دالة في سرعة السيارة. استخدم البيانات التالية في إيجاد :

x سرعة السيارة	4	3	5	8	4	3	3	4	3	5	7	3	8
y عدد العلب التالفة	63	54	86	36	65	69	28	75	53	33	168	47	52

(أ) معادلة الانحدار الخطى المقدرة.

(ب) 95% فترة ثقة لكل من β_0, β_1

-١١- قام مركز تجاري بدراسة العلاقة بين تكاليف الإعلانات الأسبوعية و المبيعات وقد تم تسجيل البيانات والجدول التالي :

x التكاليف	3	5	8	4	3	3	4	3	5	7	3	8
y المبيعات $	63	54	86	36	65	109	28	75	53	168	47	52

(أ) أوجد معادلة الانحدار الخطي المقدرة؟

(ب) أوجد تقدير للتباين σ^2؟

ج أوجد 95% فترة ثقة للمعلمة β_1 ؟

(د) أختبر فرض العدم $H_0 : \beta_1 = 0$ ضد الفرض البديل $H_1 : \beta_1 \neq 0$ وذلك عند مستوى معنوية 0.05.

-١٢- أجريت دراسة في نهر ما على عينة الثلج في الماء و تم الحصول على البيانات التالية :

x كمية الثلج في أبريل	23.1	31.8	31.7	31.0	23.0	34.1	24.1	51.1
y كمية المياه من أبريل إلى مايو	10.2	16.2	18.1	16.9	16.2	10.4	23.0	24.9

x كمية الثلج في أبريل	36.9	30.4	25.1	12.3	35.0	31.4	21.0	27.5
y كمية المياه من أبريل إلى مايو	22.7	14.0	12.8	8.7	17.3	14.8	10.4	15.1

(أ) أوجد معادلة الانحدار الخطي المقدرة ؟

(ب) أوجد تقدير للتباين σ^2 و 95% فترة ثقة للمعلمة β_0, β_1 ؟

-١٣- البيانات التالية تمثل عدد السيارات الخاصة والتي لديها رخصة سارية خلال التسع سنوات الأخيرة في بلد ما. وقد تم تسجيل البيانات من قبل شركة لبيع السيارات وذلك للتعرف على مبيعاتها من السيارات .

x السنة	1	2	3	4	5	6	7	8	9
y عدد السيارات	8247	8919	4513	10303	10816	11228	11515	112059	12717

(أ) أوجد معادلة خط الانحدار المقدرة ؟

(ب) أوجد عدد السيارات الخاصة برخصة سارية في العام ؟

-١٤- يحتوى الجدول التالي على كمية مركب كيميائي والذي يذوب في 100 جرام من الماء عند درجات حرارة مختلفة .

C^o x	x جرام		
0	7	5	7
15	11	9	13
30	24	20	23
45	30	32	25
60	43	38	41
75	47	50	43

(أ) أوجد معادلة الانحدار الخطى المقدرة ؟

(ب) أوجد كمية المركب الكيميائي والذي يذوب في 100 جرام من الماء عند $50c^o$ ؟

(ج) أختبر فرض العدم H_0 : العلاقة خطية ضد الفرض البديل H_1 : العلاقة غير خطية ؟

-١٥- قام باحث بدراسة العلاقة بين الضغط (المتغير المستقل ، Kg/mm^3)، والزمن اللازم لقطع ألواح من الصلب القاتل للصدأ (المتغير التابع ، بالساعات) في وسط ما والبيانات في الجدول التالي :

i	1	2	3	4	5	6	7	8	9	10
الضغط x_i	2.5	5	10	15	175	20	25	30	35	40
زمن القطع y_i	63	88	55	61	62	37	38	45	46	19

(أ) أوجد شكل الانتشار ؟

(ب) أوجد معادلة خط الانحدار المقدرة ؟

(ج) أختبر معنوية كل من β_0, β_1.

– ١٦– من البيانات التالية أرسم شكل الانتشار وأوجد معادلة الانحدار الخطي المقدرة .

xحجم المبيعات (بالألف دولار) لسلعة ما	5	6	7	8	9	10
y السعر (بالألف دولار)	74	77	82	86	92	95

وما توقعك عن السعر عندما يكون حجم المبيعات 7500؟

–١٧– البيانات التالية تم تسجيلها خلال 8 فترات .

الفتــرة	الوحدات المصدرة x	التكاليف الكلية y
1	10000	32000
2	20000	39000
3	30000	58000
4	40000	52000
5	50000	61000
6	60000	70000
7	7000	64000
8	90000	66000

(أ) أرسم شكل الانتشار ؟

(ب) أوجد خط الانحدار المقدر ؟

(ج) أوجد التكاليف الكلية عندمـــا يكـــون عـــدد الوحـــدات المصـــدرة , 48750 12600وحده.

–١٨– إذا كان لديك البيانات التالية :

$$n = 12 \;,\; \Sigma x_i y_i = 2000 \;,\; \bar{y} = 19 \;,\; \Sigma y_i^2 = 4612,$$
$$\bar{x} = 8 \;,\; \Sigma x_i^2 = 1502$$

(أ) أوجد معادلة الانحدار المقدرة ؟

(ب) أختبر الانحدار عند $\alpha=0.05$ ؟

–١٩– إذا كان لديك البيانات التالية :

$$n = 25 \;,\; \Sigma(y - \bar{y})^2 = 500 \;,\; \Sigma(y_i - \hat{y}) = 125,$$
$$\Sigma(\hat{y}_i - \bar{y})^2 = 375$$

أوجد تقدير للتباين σ^2 ومعامل التحديد؟

-٢٠- إذا كان لديك البيانات التالية :

$$n = 100 \;,\; \Sigma(y_i - \bar{y})^2 = 8000 \;,\; \Sigma(\hat{y}_i - \bar{y})^2 = 7000,$$

قدر معامل الارتباط مع العلم أن إشارة β_1 موجبة ؟

-٢١- أخذت عينة عشوائية من 10 موظفين وتم الحصول على البيانات التالية :

x عدد سنوات الدراسة	12.5	14	12	16	11.0	16.0	8.0	11.2	15.0	12.7
y الأجر بالآلف دولار	20.2	13.7	19.8	12.5	21.8	35.7	39.2	22.6	25.9	17.3

(أ) ارسم شكل الانتشار ؟

(ب) قدر معادلة الانحدار الخطى

(ج) أختبر فرض العدم $H_0: \beta_1 = 0$ ضد الفرض البديل $H_1: \beta_1 \neq 0$ عند مستوى معنوية $\alpha=0.05$.

-٢٢- قامت شركة بإجراء اختبار لـ 500 موظف جديد فإذا كانت x تمثل الدرجة التي حصل عليها الموظف الجديد. وبعد ثلاثة سنوات أعطيت درجات على الأداء في العمل y الدرجة التي حصل عليها . فإذا كان لديك البيانات التالية :

$$\bar{x} = 100 \quad , \quad s_x = 10 \quad , \quad \bar{y} = 130 \quad , \quad s_y = 20 \quad , \quad r = 0.7$$

حيث S_x , S_y تمثل الانحراف المعياري من العينة لقيم x , y على التوالي :

(أ) أوجد معادلة الانحدار المقدرة ؟

(ب) أوجد قيمة y عندما $x = 90$ و $x = 125$ ؟

-٢٣- قام باحث بالحصول على بيانات من 1000 مريض نفسي وذلك لمدة 5 سنوات ، فإذا كان x يمثل الدرجة التي حصل عليها الشخص في بداية العلاج ، y الدرجة التي حصل عليها بعد تلقى العلاج باستخدام البيانات التالية :

$$\Sigma x_i = 3000 \quad , \quad \Sigma y_i = 5000 \quad , \quad \Sigma x_i y_i = 3000$$
$$\Sigma x_i^2 = 14000.$$

(أ) أوجد معادلة الانحدار المقدرة؟

(ب) أوجد قيمة y عندما $x=4$؟

(ت) إذا كانت قيمة $s_y=10$ أوجد معامل الارتباط r.

-٢٤- للبيانات التالية :

$$\bar{x} = 200 \ , \ \bar{y} = 90 \ , \ r = -0.9 \ , \ s_x = 9 \ , \ s_y = 5$$

أوجد معادلة الانحدار المقدرة وتنبأ بقيمة y عندما **x = 164** .

-٢٥- في مكتب للشرطة تم إجراء دراسة للعلاقة بين عدد الجرائم في اليوم وأعلى درجة حرارة في اليوم . اختبرت عينة عشوائية في **10** أيام وتم الحصول على البيانات التالية .

x أعلى درجة حرارة	12	89	40	52	75	60	50	20	32	90
y عدد الجرائم	2	15	6	8	14	12	7	3	5	16

(أ) أوجد معادلة الانحدار المقدرة؟

(ب) أختبر فرض العدم $H_0: \beta_1 = 0$ ضد الفرض البديل $H_1 : \beta_1 \neq 0$ عند مستوى معنوية $\alpha=0.05$.

-٢٦- الجدول التالي يعطى العمر لأحد النباتات (بالأسابيع) وطوله بالسنتيمتر .

العمر بالأسبوع x	1	2	3	4	5	6	7
الطول بالسنتيمتر y	5	13	16	33	23	38	40

أوجد معادلة خط الانحدار المقدرة ؟

-٢٧-أجريت تجربة لتقدير العلاقة بين درجة الحرارة ومعدل دقات القلب في الضفدعة المسماة **Rana Pipiens** والبيانات معطاة في الجدول التالي :

الحيوان	1	2	3	4	5	6	7	8	9
x درجة الحرارة	2	4	6	8	10	12	14	16	18
y دقات القلب بالدقيقة	5	11	11	14	22	23	32	29	32

(أ)أوجد معادلة الانحدار المقدرة؟

(ب) أختبر فرض العدم $H_0: \beta_1 = 0$ ضد الفرض البديل $H_1 : \beta_1 \neq 0$ عند مستوى

معنوية $\alpha=0.05$.

-٢٨- أجريت تجربة لتقدير العلاقة بين العمر ودقات القلب (في الدقيقة) في الإنـــاث الذيـــن أعمارهم تتراوح من واحد إلى 15سنة . استخدم البيانات المعطاة في الجدول التالي في إيجاد :

(أ) معادلة الانحدار المقدرة؟

(ب) أختبر معنوية معادلة الانحدار؟

(ج) أوجد 95% فترة ثقة لكل من B_0 , B_1 .

الأنثى	1	2	3	4	5	6	7	8	9	10
العمر x	1	2	3	4	5	6	7	8	9	10
دقـــات القلب y	111	108	108	102	99	92	93	88	90	90
الأنثى	11	12	13	14	15					
العمر x	11	12	13	14	15					
دقـــات القلب y	88	84	83	83	82					

-٢٩- يعطى الجدول التالي عدد سنوات الخبرة للأستاذ في جامعة ما والدخل السنوي بـــالدولار (بالآلاف):

السنوات x	5	10	15	20	25	30
الدخل y	59.0	69.0	78.0	88.0	97.5	107.0

(أ) أختبر العلاقة بين سنوات الخبرة والدخل السنوي بيانيا؟

(ب) اقترح شكل العلاقة بين المتغيرين وقدر معالم النموذج؟

-٣٠- يعطى الجدول التالي أعمار الزوج والزوجة بالسنوات لعينة من 6 أزواج:

x عمر الزوجة	35	25	51	25	53	42
y عمر الزوج	38	25	49	31	55	44

أوجد معامل الارتباط .

-٣١- قام باحث بدراسة العلاقة بين ضغط الغاز (مقـــاس milliliters) ودرجـــة الحـــرارة (مقاسه k^{o}) وتم الحصول على البيانات التالية :

x درجة الحرارة	200	250	300	350	200	250	300	350

←

y الضغط	423	362	302	241	440	374	315	251

ارسم شكل الانتشار وهل تعتقد أن النموذج الخطى مناسب لتوفيق البيانات السابقة ؟

-٣٢- الجدول التالي يمثل درجات امتحان أعمال السنة ودرجات الامتحـــان النـــهائي لعينـــة عشوائية من 8 طلبة .

x درجة أعمال السنة	98	95	93	80	71	70	49	75
y درجة الامتحان النهائي	98	98	84	40	33	77	65	80

(أ) أوجد معادلة الانحدار المقدرة؟

(ب) أوجد الدرجة النهائية لطالب إذا كانت درجته في أعمال السنة 83.

-٣٣- تعمل آلة عند سر عات مختلفة ولكن السرعة العالية تؤثر على عمر يد المثقاب البيانـــات التالية تمثل أعمار يد المثقاب عند دورات مختلفة من الآلة.

xعدد الدورات في الدقيقة	42	41	40	32	32	31	28	26	26	23	20	43	20	18
y عمر يـــد المثقـــاب	109	105	95	97	106	125	129	140	138	162	171	69	54	162

(أ) ارسم شكل الانتشار ؟

(ب) قدر عمر يد المثقاب عندما تكون عدد دورات الآلة 30 دورة في الدقيقة ؟

-٣٤- أعطى اختبار في المعلومات العامة يتكون من 100 سؤال إلى 20 طالب في أعمار مختلفـــة وكانت النتائج كالتالي :

عدد الإجابات الصحيحة	العمر		الطالب
	شهر	سنوات	
40	8	16	A
45	2	16	B
47	9	17	C
46	1	17	D
67	6	18	E
45	7	18	F
53	10	19	G
54	4	20	H

المطلوب (أ) رسم شكل الانتشار ؟

(ب)إيجاد معادلة الانحدار الخطى المقدرة ؟

(ج) أوجد باستخدام خط الانحدار المقدر عدد الإجابات الصحيحة إذا كان عمر الطالب 17 سنة .

-٣٥- أجرى ثلاثة باحثين في الجيولوجيا في القطب الشمالي بحث أستمر خمسة فصول شتوية وذلك لدراسة العلاقة بين عدد الأيام التي تنخفض فيها درجة الحرارة عن 50^oF وطول القرون لحيوان الموظ (حيوان ضخم من الأيائل). أوجد معامل الارتباط الذي يعتمد على هذه البيانات .

x عدد الأيام التي تنخفض درجة الحرارة عن 50^oF	10	30	10	20	30
متوسط الطول لفروق الحيوان (بالمتر)	0.8	1.0	0.5	0.8	0.9

-٣٦- أحسب معامل الارتباط الخطى بين الفئات التالية من الأرقام

درجات اختبار ذكاء	103	86	116	112	97	109	126	93	105	84	121	97
الدخل الأسبوعي عند العمر 23	21	20	9	14	13	22	17	18	18	19	16	20

-٣٧- الجدول الآتي يبين طول الجمجمة X وعرضها Y بالمليمتر والمطلوب إيجاد معامل الارتباط الخطي .

x الطول	71	58	72	73	79	66	76	70	80	63
y العرض	35	39	37	42	46	39	38	45	42	40

-٣٨- الجدول التالي يبين المبيعات اليوم بالجنية X لعشرة عمال في متجر ومدة خدمتهم Yبالسنين أوجد معامل الارتباط الخطى .

x	5	4	5	6	9	10	9	12	10	9
y	5	2	4	9	8	4	10	11	10	7

-٣٩- الجدول التالي يعرض عدد العدسات التي تنتجها إحدى المصانع وتكلفة العدسة الواحدة بالدولار .

xعدد العدسات	12	10	5	3	1
yتكلفة العدسة	5	7	10	15	20

أوجـــد معامل الارتباط؟

٤٠-اختيرت عينة عشوائية من 12 ورقة من شجرة ما وتم قياس الطول والعرض لكل ورقـــة إلى أقرب ملليمتر . البيانات معطاة في الجدول التالي :

الورقة	1	2	3	4	5	6	7	8	9	10	11	12
العرض	35	21	25	35	26	40	35	40	35	42	23	25
الطول	55	44	46	60	55	57	64	68	5	61	46	44

(أ) أوجد معامل الارتباط ؟

(ب) أختبر معنوية معامل الارتباط المجتمع ρ؟

أوجـــد معامل الارتباط؟

(ت) أختبر معنوية معامل الارتباط المجتمع ρ؟

٤١- الجدول التالي يبين درجات مجموعة مكونة من 10 طلاب في كل من مـــادتي الإحصـــاء والرياضيات في أحد الامتحانات لأعمال الفصليين .

x الإحصاء	17	13	8	17	14	10	7	17	8	9
y الرياضيات	18	14	6	16	14	9	10	13	7	8

(أ) أوجد معامل الارتباط ؟

(ب) أوجد 95% فترة ثقة للمعلمة ρ ؟

٤٢- الجدول التالي يعطى أطوال الأب X وأطوال ابنه الأصغر عند بلوغه سن معينY (البيانات مقاسه لاقرب بوصة).

x	68	64	70	72	69	74
y	67	68	69	73	66	70

أوجد معامل الارتباط الخطي البسيط ؟

٤٣-لدراسة العلاقة بين عدد الكيلوجرامات التي يفقدها شخص في برنامج لإنقاص الـــوزن وعدد الأسابيع التي يقضيها لإنقاص الوزن ، اختيرت عينة عشوائية من 5 أشخاص ممن يتبعـــون هذا البرنامج الغذائي وتم الحصول علي البيانات التالية :

x	3	2	1	4	5
y	6	5	4	9	11

أوجد معامل الارتباط الخطي ؟

٤٤- لدراسة العلاقة بين المدة اللازمة لتدريب العامل في مصنع ما وعمر العامل تم الحصـــول علي البيانات التالية :

x العمر (بالسنوات)	18	20	21	27	23	34	24	42	38	44
y المدة اللازمة للتدريب	8	5	6	8	7	11	8	10	6	8

أحسب معامل الارتباط الخطي البسيط ؟

-٤٥- البيانات التالية تعطي متوسط درجة الحرارة اليومية وكمية الكهرباء المستهلكة وذلك خلال ثمانية أيام اختيرت في سنة ما :

x الحرارة (F^0)	37	32	35	40	40	44	42	48
y استهلاك الكهرباء Megawatt Hours	3.7	3.8	3.7	3.6	3.7	3.4	3.4	3.3

أحسب معامل الارتباط الخطي البسيط ؟

-٤٦- لدراسة العلاقة بين عدد الموظفين في شركة للخدمات وعدد الطلبات المتدفقة على الشركة تم الحصول على البيانات التالية :

عدد الموظفين	1	2	3	5	8	12	15
عدد الطلبات	10	15	20	25	30	35	40

(أ) أوجد نموذج الانحدار الأسي المقدر ؟

(ب) أوجد عدد الموظفين عندما تكون عدد الطلبات 45 ؟

-٤٧-لدراسة العلاقة بين النسبة المئوية للعاطلين عن العمل والتغير في الأجور خلال عشر سنوات تم الحصول على البيانات التالية :

النسبة % x العاطلين عن العمل	1 11.6	2 12.2	3 12.5	4 11.7	5 11.6	6 12.1	7 12.6	8 11.7	9 11.5	10 11.6
y المتغير في الأجور	5.0	3.2	2.7	2.1	4.1	2.7	2.9	4.6	3.5	4.4

أوجد معامل الارتباط الخطي البسيط ؟

٤٨- أجريت تجربة لاختبار تأثير نوع معين من العقاقير على سريان الادرينا لين في حيوان مـا. إذا كان Y يمثل عدد الملليجرامات من العقار اللازمة لإحداث سريان من الأدرينالين مقـداره X وإذا كانت العلاقة التي تربط بين X,Y يمثلها معادلة على الشكل $\mu_{Y|x} = b\,a^x$ حيث a , b ثابتان. أوجد تقديرات لمعلمتين a , b إذا كان لديك البيانات التالية :

x	1	2	3	4	5	6
y	3.1	6.2	11.3	22.0	48.0	92.0

٤٩- أجريت تجربة على أحد الأنواع الجديدة من السيارات لتحديد مسافة التوقف للسـيارة عند سرعات مختلفة ، وكانت نتائج التجربة كما يلي :

x السرعة بالميل/ساعة	20	30	40	50	60	70
y مسافة التوقف بالقدم	50	85	130	200	280	290

إذا كانت العلاقة بين السرعة ومسافة التوقف هي على الشكل :

$$\mu_{Y|x} = a_0 x^{b_0}$$

استخدم البيانات السابقة في تقدير معالم النموذج ثم قدر مسافة التوقف لسيارة سرعتها 53 ميلا في الساعة .

٥٠- البيانات التالية تمثل أعداد السكان بالمليون في مدينة ما خلال الفترة من 1980-1972.

فإذا كانت العلاقة بين السنوات وأعداد السكان على الشكل :

$$\mu_{Y|x} = a_0 x^{b_0}$$

أوجد معادلة الانحدار المقدرة ؟

x العام	1970	1971	1972	1973	1974	1975
y عدد السكان بالمليون	1.010	1.050	1.060	1.080	1.111	1.156
x العام	1976	1977	1978	1979	1980	
y عدد السكان بالمليون	1.155	1.170	1.201	1.230	1.330	

٥١- في دراسة أجريت على إحدى أنواع الثدييات وجد أن حجم المخ يتغير مع وزن الجسم من فرد لآخر وأن العلاقة بين حجم المخ ووزن الجسم على الشكل :

$$\mu_{Y|x} = a_0 x^{b_0}$$

استخدم البيانات التالية لإيجاد معادلة الانحدار المقدرة ؟

x وزن الجسم	30	35	37	40	41	44	46	47	49	52	54
y حجم المخ	360	379	380	390	409	408	412	419	425	435	439

٥٢- الجدول التالي يوضح متوسط طول الجمجمة مقاساً بالملليمترات ومتوسط طول الوجه y مقاساً بالملليمترات لنوع معين من القردة. فإذا كانت العلاقة بين طول الجمجمة وطول الوجه على الشكل

$$\mu_{Y|x} = a_0 x^{b_0}$$

استخدم البيانات التالية للحصول على تقديرات لكل من a_0, b_0 .

x طول الجمجمة	75.5	99.1	107.8	113.7	115.25	120.0
y طول الوجه	30	63.5	93.7	131.0	141.2	143.22

٥٣- البيانات التالية تمثل سعر بيع السيارة وعمر السيارة لماركة خاصة بالسنوات :

x العمر بالسنوات	1	2	2	5	5	5
y سعر البيع	2350	1690	1749	1390	981	890

إذا كانت العلاقة بين عمر السيارة وسعر البيع على الشكل : $\mu_{Y|x} = \gamma\ \delta^{x}$

(أ) أوجد تقديرات لكل من δ , γ.

(ب) أوجد سعر البيع عندما يكون عمر السيارة 4 سنوات .

٥٤- يعطي الجدول التالي الضغط لغاز ما عند قيم مختلفة من الحجم .

x الحجم $in.^3$	50	60	70	90	100

←

الضغط ib/in^3 y	63.6	50.2	40.1	24.8	7.7

فإذا كان قانون الغاز معطي بالعلاقة $\mu_{Y|x} \quad x^a = c$ حيث a , c ثوابت أوجد :

(أ) تقديرات المربعات الصغرى للمعالم a , c

(ب) أوجد y عندما x = 70 .

-٥٥- في دراسة عن العلاقة بين عدد الحوادث في مدينة ما خلال 10 سنوات وبين عدد الإشارات تم الحصول علي البيانات التالية (من 1969 إلي 1978).

السنوات	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
x عدد الإشارات	452	425	500	517	350	480	603	611	650	718
y عدد الحوادث	42	1	6	5	1	4	7	10	15	18

استخدم النموذج الأسى في حساب معادلة الانحدار المقدرة ؟

-٥٦-البيانات التالية تعطى عدد العمال الزراعيين في الفترة من 1968 إلي 1975 .

السنوات	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975
عدد العمال	1050	1000	900	800	750	675	500	450	425	350	300

(أ) أوجد معادلة الانحدار المقدرة تحت فرض النموذج الأسى ؟

(ب) قدر معادلة الانحدار من الدرجة الثانية ؟

(ج) قدر معادلة الانحدار الخطية ؟

(د) أى المعادلات منطقية ولماذا ؟

-٥٧- إذا كان لديك البيانات التالية :

x	0	1	2	3	4	5	6	7	8	9
y	9.0	7.2	3.1	4.5	4.7	2.8	5.6	7.0	8.7	10.1

(أ) المطلوب توفيق هذه البيانات باستخدام النموذج :

$$\mu_{Y|x_1,x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

(ب) أوجد y عند x = 2 .

-٥٨- الجدول التالي يعطي أعداد البيتزا المباعة ومتوسط التكلفة في محل لبيع البيتزا .

x عدد الوحدات المباعة	100	110	120	130	140	150	160	170	180
y متوسط التكلفة لوحده البيتزا	0.90	0.88	0.85	0.72	0.65	0.70	0.73	0.79	0.74

(أ) ارسم شكل الانتشار ؟

(ب) أوجد معادلة الانحدار المقدرة تحت فرض معادلة من الدرجة الثانية ؟

-٥٩- يعطى الجدول التالي عدد القوارب المباعة في مدينة ما خلال السنوات من 1970 إلي 1975 .

x السنوات	1970	1971	1972	1973	1974	1975
y عدد القوارب المباعة	100	90	85	90	75	80
x السنوات	1976	1977	1975			
y عدد القوارب المباعة	86	110	115			

أستخدم معادلة من الدرجة الثانية لتوفيق هذه البيانات وأوجد عدد القوارب المباعة سنة 1981

-٦٠- استخدم البيانات التالية في إيجاد نموذج الانحدار الخطي المتعدد المقدر ؟

x_1 السنوات	1	2	3	4	5	6	7	8	9	10
x_2 عدد الإشارات	452	425	500	517	350	480	603	611	850	718
y عدد الحوادث	42	1	6	5	1	4	7	10	15	18

-٦١- يتأثر محصول الفراولة بكمية الإمطار وكمية السماد المستخدم أستخدم البيانات في الجدول التالي لتوفيق معادلة انحدار خطي متعدد باستخدام كمية الإفطار وكمية السماد كمتغيرات مستقلة ؟

x_1 كمية الامطار	16	22	23	13	17	25	18	20	21	19	22
x_2 السماد بالطن	510	450	500	425	450	475	515	500	490	510	525

←

المحصول y	1000	450	1200	700	800	1100	1050	1150	1000	950	1300

-٦٢- تبيع شركة لإنتاج السجق منتجاها من خلال بعض المراكز التجارية البيانات التالية تعطي مبيعات الشركة (بالألف دولار) في 10 مراكز وعدد الوحدات المباعة (بالمائة) وعدد المراكز التجارية أوجد معادلة الانحدار الخطي المتعدد المقدرة ؟

الموقع	1	2	3	4	5	6	7	8	9	10
المبيعات y	75	38	25	98	93	54	78	85	65	88
عدد x_1 الوحدات المباعة	15	10	7	21	14	8	14	24	9	23
عدد x_2 المراكز التجارية	25	29	15	25	11	13	22	14	13	11

-٦٣- للبيانات التالية :

x_1	1	4	9	11	3	8	5	10	2	7	6
x_2	8	2	-8	-10	6	06	0	-12	4	-2	-4
y	6	8	1	0	5	3	2	-4	10	-3	5

استخدم النموذج التالي :

$$\mu_{Y|x_1,x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

في إيجاد

(أ) معادلة الانحدار الخطي المتعدد المقدرة ؟

(ب) جدول تحليل التباين واستخدام $\alpha = 0.05$ في اختبار معنوية b_0, b_1, b_2 .

(ج) معامل التحديد ؟

-٦٤- البيانات التالية تعطي بيانات لثلاثة متغيرات X, Y, Z حيث X , Y متغيرات مستقلة :

$$\sum_{i=1}^{20} x_i = 23,\quad \sum_{i=1}^{20} y_i = 40, \sum_{i=1}^{20} z_i = 67,\quad \sum_{i=1}^{20} x_i^2 = 105,\quad \sum_{i=1}^{20} y_i^2 = 294,$$

$$\sum_{i=1}^{20} x_i y_i = \sum_{i=1}^{20} x_i z_i = 290,\quad \sum_{i=1}^{20} y_i z_i = 4888, \sum_{i=1}^{20} z_i^2 = 815$$

استخدم البيانات السابقة في إيجاد معادلة الانحدار الخطي المقدرة ؟

-٦٥- تتوقف قيمة المبيعات داخل أي جمعية استهلاكية على المربع السكنى الذي توجـد بـه الجمعية وذلك من حيث عدد السكان ومتوسط الدخل الشهري لقاطن هذا الحي بفرض أن كل جمعية يشتري منها سكان الحي الذي تتواجد به وفيما يلي بيانات لعشر جمعيات استهلاكية .

الجمعية	1	2	3	4	5	6	7	8	9	10	11	12
x_1 عدد السكان	275	175	375	200	175	260	90	300	190	50	400	270
x_2 متوسط دخل العمل	240	320	370	282	320	375	300	240	200	210	250	400
المبيعات بالآلاف جنية	160	115	220	130	119	165	80	190	115	50	250	144

أوجد معادلة الانحدار الخطي المتعدد المقدره ؟

-٦٦- في دراسة عن العلاقة بين امتصاص الماء في دقيق القمح والخواص المختلفة للدقيق وتحـت فرض نموذج انحدار خطي متعدد تم الحصول على البيانات في الجدول التالي حيث Y(%) تمثـل كمية امتصاص الماء و $x_1(\%)$ كمية البروتين و $x_2(\%)$ كمية النشا الذي يتعرض للفقد التحطم مقاس بوحدات Farrand أوجد معادلة الانحدار المقدرة ؟

x_1	8.5	8.9	10.6	10.2	9.8	10.8	11.6	12.0	12.5	10.4
x_2	2	3	3	20	22	20	31	32	31	28
y	30.9	32.7	36.7	41.9	40.9	42.9	46.3	47.2	44.0	47.7
x_1	12.2	11.9	11.3	13.0	12.9	12.0	12.9	13.1	11.4	13.2
x_2	36	28	30	27	24	25	28	28	32	28
y	43.9	46.8	46.2	47.0	46.8	45.9	48.8	46.2	47.8	49.2

الفصل الحادي عشر

تحليل التباين

Analysis of Variance

(١١-١) مقدمـــــة Introduction

ذكرنا في الفصل التاسع اختبار t والذي يخص الفرق بين متوسطي مجتمعين وذلك تحت شروط معينه. في كثير من الأحيان يحتاج الباحث إلى مقارنة متوسطات ثلاثة مجتمعات فأكثر. فعلي سبيل المثال إذا كان لدينا أربع طرق للتعليم A, B, C, D يحوي الواحد منها كـــل الأطفال الذين يتلقون تعليمهم بإحدى هذه الطرق والمطلوب مقارنـــة متوســـطات المعرفـــة المكتسبة في كل من الطرق المختلفة. يمكن استخدام اختبار t لمقارنة متوسطي مجتمعين لكل زوج من المجتمعات الأربعة ، أى استخدام اختبار t لمقارنـــة الطريقـــة A بالطريقـــة B ثم استخدامه مرة أخري لمقارنة الطريقة A بالطريقة C وهكذا ، إلا أن هذه الطريقة لها مشاكل كثيرة منها:

(أ) غير عملية حيث يزداد عدد المقارنات بسرعة كلما زاد عدد المجتمعات فمثلا في المثـــال السابق نحتاج لإجراء اختبار t ستة مرات لأن $\binom{4}{2}=\frac{4!}{2!2!}=6$. بصورة عامـــه عدد المقارنات الزوجية لعدد k من المتوسطات يساوى $\binom{k}{2}=\frac{k!}{2!(k-2)!}$.

(ب) زيادة احتمال الوقوع في خطأ من النوع الأول أى رفض فرض العـــدم وهـــو صحيح وذلك لأن عدد المقارنات الزوجية ومستوى المعنويـــة يرتبطـــان باحتمـــال الوقوع في خطأ من النوع الأول من خلال العلاقة التالية : $1-(1-\alpha)^k$ حيـــث k هي عدد المقارنات الزوجية و α مستوى المعنوية والذي سوف يحدد عنـــد إجـــراء مقارنة واحدة فقط . وعلى ذلك إذا كانت $k=6$ ومســـتوى المعنويـــة $\alpha=0.05$ ، والذي يحدد لكل مقارنة زوجية ، فإن احتمال الوقوع في خطأ من النوع الأول هو :

$$1-(1-\alpha)^k=1-0.95^6=1-0.73509=0.26491.$$

أى ما يقرب من خمسة أمثال مستوى المعنوية $\alpha=0.05$ والذي سوف يحـــدد عنـــد مقارنـــة واحدة فقط للمتوسطات الستة (الأربعة) في آن واحد. لحسن الحظ فإنه يمكن التغلب علـــى المشـــاكل السابقة ، ومشاكل أخري ، باستخدام اختبار إحصائي يسمى تحليل التباين والـــذي يعتـــبر واحد من أكثر الطرق الإحصائية استخداما . سوف نوضح أسلوب تحليل التبـــاين بالمثـــال التالي. إذا أجريت تجربة زراعية لدراسة تأثير الأوقات المختلفة للزراعة (فبراير – مـــارس – نوفمبر – أكتوبر) على إنتاجيه محصول القصب وإذا كان اهتمامنا هو اختبار فرض العدم أن متوسط إنتاجية محصول القصب واحد للأوقات المختلفة. يعتمد أسلوب تحليل التباين، في هذه الحالة، على تجزئة الاختلاف الكلي للمشاهدات إلى مكونين لهما معني يستخدمان في قيـــاس

المصادر المختلفة للاختلاف . المكون الأول يقيس الاختلاف الذي يرجع إلى خطأ التجربة والثاني يقيس الاختلاف الذي يرجع إلى خطأ التجربة بالإضافة إلى الاختلاف الذي يرجع إلى أوقات الزراعة الأربعة . عندما يكون فرض العدم صحيح ، أي أن متوسط إنتاجية محصول القصب واحدة للأوقات المختلفة ، فإن كلا من المكونين سوف يمدوننا بتقديرين مستقلين لخطأ التجربة ، وعلى ذلك يعتمد اختبارنا على المقارنة بين المكونين باستخدام توزيع F .

بفرض أن اهتمامنا سوف يكون في مقارنة متوسط إنتاجية محصول القصب عند أوقات مختلفة للزراعة وباستخدام ثلاثة طرق للزراعة (1,2,3). اهتمامنا في هذه الحالة سوف يكون في اختبار ما إذا كان الاختلاف في إنتاجية محصول القصب يرجع إلى الفروق في مواعيد الزراعة أو الفروق في طرق الزراعة أو ربما الفروق في كلاهما. يعتمد تحليل التباين ، في هذه الحالة ، على تجزئة الاختلاف الكلي لإنتاجية محصول القصب إلى ثلاثة مكونات ، الأول يقيس خطأ التجربة فقط والثاني يقيس خطأ التجربة بالإضافة إلى أي اختلاف يرجع إلى مواعيد الزراعة المختلفة ، والثالث يقيس خطأ التجربة بالإضافة إلى أي اختلاف يرجع إلى طرق الزراعة المختلفة . وعلى ذلك فإن مقارنة المكون الأول بالثاني سوف يمدنا باختبار الفرض أن متوسط إنتاجية محصول القصب واحدة عند مواعيد الزراعة المختلفة. بنفس الشكل يمكن اختبار الفرض أن متوسط إنتاجية محصول القصب واحدة لطرق الزراعة المختلفة عن طريق مقارنة المكون الأول بالثالث .

إذا صنفت المشاهدات وفقاً لصفة (خاصية) واحدة مثل الاختلاف في طرق الزراعة أو الجنس أو العمر ... الخ فسوف يكون لدينا تصنيف أحادي **one-way classification** . أما إذا صنفت المشاهدات وفقا لصفتين مثل أصناف القمح وأنواع الأسمدة فسوف يكون لدينا تصنيف ثنائي **two-way classification** . في البنود التالية سوف نتناول طرق تحليل التباين في كلا التصنيفين .

(١١-١) التصنيف الأحادي One-way Classification

بفرض أن عينات عشوائية من الحجم n تم اختيارها من k من المجتمعات . سوف نفترض أن المجتمعات التي عددها k مستقلة وتتبع توزيعات طبيعية بمتوسطات $\mu_1,\mu_2,\dots,\mu_k$ وتباين مشترك σ^2 . المطلوب اختبار فرض العدم :

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

ضد الفرض البديل :

H_1 : واحد على الأقل من μ_i يختلف عن الباقي

بفرض أن x_{ij} ترمز للمشاهدة رقم j المختارة من المجتمع رقم i وأن المشاهدات تم ترتيبها في جدول (١١-١) حيث $T_{i.}$ ترمز لمجموع كل المشاهدات في العينة المختارة من المجتمع رقم i و $\bar{x}_{i.}$ ترمز لمتوسط كل المشاهدات في العينة المختارة من المجتمع رقم i و $T_{..}$ ترمز لمجموع كل المشاهدات التي عددها nk و $\bar{x}_{..}$ ترمز لمتوسط كل المشاهدات التي عددها nk .

جدول (١١-١)

	المجتمعات	
	1 2 ... i ... k	
	x_{11} x_{21}.... x_{i1}... x_{k1} x_{12} x_{22}.... x_{i2}... x_{k2} ⋮ ⋮ ⋮ ⋮ x_{1n} x_{2n}.... x_{in}... x_{kn}	
المجموع	$T_{1.}$ $T_{2.}$... $T_{i.}$... $T_{k.}$	$T_{..}$
المتوسط	$\bar{x}_{1.}$ $\bar{x}_{2.}$... $\bar{x}_{i.}$... $\bar{x}_{k.}$	$\bar{x}_{..}$

يمكن التعبير عن كل مشاهدة وفقا للنموذج الرياضي التالي :

$$x_{ij} = \mu_i + \epsilon_{ij},$$

حيث ϵ_{ij} يقيس انحراف المشاهدة رقم j في العينة رقم i عن متوسط المجتمع رقم i

وبوضع $\mu_i = \mu + \alpha_i$ حيث :

$$\mu = \frac{\sum_{i=1}^{k} \mu_i}{k}$$

فإنه يمكن كتابة النموذج أعلاه على الشكل :

$$x_{ij} = \mu + \alpha_i + \epsilon_{ij},$$

تحت شرط أن $\sum_{i=1}^{k} \alpha_i = 0$ حيث α_i تعبر عن تأثير المجتمع رقم i . وباستعمال النموذج الأخير يصبح فرض العدم $H_0 : \mu_1 = \mu_2 = ... = \mu_k$ مكافئ للفرض :

$$H_0 : \alpha_1 = \alpha_2 = ... = \alpha_k = 0$$

ضد الفرض البديل :

H_1: واحد على الأقل من α_i لا يساوى صفراً

اختبارنا سوف يعتمد على مقارنة تقديرين مستقلين لتباين المجتمع σ^2. يتم الحصول على التقديرين بتجزئه الاختلاف الكلي للمشاهدات إلى مكونين . من المعروف أن التباين لكل المشاهدات مجتمعه في عينة واحدة من الحجم nk يعطى من الصيغة :

$$s^2 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{..})^2}{nk - 1}$$

البسط في الصيغة السابقة يسمى مجموع المربعات الكلـــي **total sum of squares** والذي يقيس الاختلاف الكلي للمشاهدات حيث :

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{..})^2 = n\sum_{i=1}^{k}(\bar{x}_{i.} - \bar{x}_{..})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{i.})^2.$$

ويمكن التعبير عن الحدود في العلاقة السابقة باستخدام الرموز كالتالي :

$$SSTO = SSC + SSE$$

حيث مجموع المربعات الكلي هو :

$$SSTO = \sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{..})^2,$$

ومجموع المربعات لمتوسطات الأعمدة **sum of squares for columns means** هو:

$$SSC = n\sum_{i=1}^{k}(\bar{x}_{i.} - \bar{x}_{..})^2,$$

ومجموع المربعات للخطأ **error sum of squares** هو :

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{i.})^2,$$

أيضا تجزئ درجات الحرية الكلية كما يلي :

nk-1= k-1 + k (n-1).

عادة يشار لمجموع المربعات لمتوسطات الأعمدة من قبل كثير من المؤلفين بمجموع المربعـــات للمعالجات **treatment sum of squares** . وهذه التسمية ترجـــع إلى الحقيقـــة أن **k** من المجتمعات المختلفة غالباً ما تصنف تبعاً لمعالجات مختلفة وعلى ذلك فإن المشاهدات x_{ij} ;**(j= 1,2,...,n)** تمثل **n** من المشاهدات المقابلة للمعالجة رقم **i** . الآن كلمة معالجـــة

تستخدم أكثر لتوضيح التصنيفات المختلفة سواء أسمدة مختلفة أو مصانع مختلفة أو منـــاطق مختلفة في مدينة ما أو محللين مختلفين .

التقدير الأول للمعلمة σ^2 ، يعتمد علي k-1 درجات حرية ، ويعطي من الصيغة :

$$MSC = \frac{SSC}{k-1}.$$

عندما يكون H_0 صحيح ، فإن MSC سوف يكون تقدير غير متحيز للمعلمة σ^2 .

التقدير الثاني المستقل للمعلمة σ^2 يعتمد علي k(n-1) درجات حرية ويعطي من الصيغة :

$$MSE = \frac{SSE}{k(n-1)}.$$

يعتبر التقدير MSE غير متحيز بصرف ، النظر عن صحة أو عدم صحة فرض العـــدم .

نعرف مما سبق أن التباين لكل مشاهدات العينة ، بدرجات حرية nk-1 ، هو :

$$s^2 = \frac{SSTO}{nk-1},$$

والذي يعتبر تقدير غير متحيز للمعلمة σ^2 عندما H_0 صحيح. النسبة :

$$f = \frac{MSC}{MSE}$$

هي قيمة لمتغير عشوائي F يتبع توزيع F بدرجات حرية $v_1 = k-1, v_2 = k(n-1)$ عندما H_0 صحيح. لمستوى معنويـــة α منطقـــة الرفـــض $F > f_\alpha(v_1, v_2)$ حيـــث $f_\alpha(v_1, v_2)$ تستخرج من جدول توزيع F في ملحق (٦) عند $\alpha = 0.05$ أو في ملحـــق (٧) عند $\alpha = 0.01$. إذا وقعت f في منطقة الرفض نرفض H_0 .

عمليا يتم أولا حساب SSTO , SSC ثم نحصل علـــي SSE بطـــرح SSC مـــن SSTO أى أن :

$$SSE = SSTO - SSC.$$

بإمكاننا حساب الصيغ السابقة والمعرفة لكل من SSTO و SSC بطريقة حسابية مبسطة (مناسبة للآلة الحاسبة) علي النحو التالي :

$$SSTO = \sum_{i=1}^{k} \sum_{j=1}^{n} x_{ij}^2 - CF,$$

حيث $CF = \frac{T^2}{nk}$ يسمي معامل التصحيح correction factor. أيضا :

$$SSC = \frac{\sum_{i=1}^{k} T_{i.}^2}{n} - CF .$$

عادةً الحسابات في تحليل التباين تلخص في جدول يسمي جدول تحليل التبـاين Analysis of Variance (عادة يسمى ANOVA) والموضح في جدول (١١-٢)

جدول (١١-٢)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الأعمدة	k-1	SSC	$MSC = \frac{SSC}{k-1}$	$\frac{MSC}{MSE}$
الخطأ	k(n-1)	SSE	$MSE = \frac{SSE}{k(n-1)}$	
الكلي	nk-1	SSTO		

مثال (١١-١) البيانات في جدول (١١-٣) تمثل الطول (مقاس بالسنتيمتر) لنباتات تم زراعتها في ثلاثة أوساط مختلفة A,B,C (5 نباتات في كـل وسط). أوجـد جدول تحليل التباين وأختبر فرض العدم أن $\mu_3 = \mu_2 = \mu_1$ وذلك عند مستوي معنويـة $\alpha=0.05$.

جدول (١١-٣)

الأوساط						
	A	10	14	18	15	12
	B	16	18	22	18	15
	C	15	12	8	10	13

الحل . المطلوب اختبار فرض العدم :

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

ضد الفرض البديل :

H_1 : واحد علي الأقل من μ_i يختلف عن الباقي

$$\alpha=0.05 .$$

$f_{.05}(2,12) = 3.89$ والمستخرجة من جدول توزيع F في ملحق (٦) عند درجات حريـة $v_1 = 2, v_2 = 12$. منطقة الرفض $F>3.89$.

$$SSTO = \sum_{i=1}^{k} \sum_{j=1}^{n} x_{ij}^2 - CF$$

$$=10^2+14^2+...+10^2+13^2-\frac{(216)^2}{15}$$
$$=3304-3110.4=193.6,$$

$$SSC=\frac{\sum_{i=1}^{k}T_i^2}{n}-CF$$
$$=\frac{69^2+89^2+58^2}{5}-\frac{(216)^2}{15}$$
$$=3209.2-3110.4=98.8.$$

تلخص النتائج في جدول تحليل التباين [جدول (١١-٤)].

جدول (١١-٤)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الأعمدة	2	98.8	49.4	6.25316*
الخطأ	12	94.8	7.9	
الكلي	14	193.6		

وبما أن f (6.25316) تقع في منطقة الرفض فإننا نرفض H_0 ونعتبر أن هناك فروقاً معنوية بين متوسطات الأوساط المختلفة. النجمة * تعني أن الفرق معنوي عند $\alpha=0.05$.

الآن بفرض أن العينات التي عددها k ذات أحجام $n_1, n_2, ..., n_K$ (عدم تساوى حجوم العينات) حيث $N=\sum_{i=1}^{k}n_i$. الصيغ المستخدمة لحساب SSTO , SSC تعطى كالآتي :

$$SSTO=\sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij}^2-CF,$$
$$SSC=\sum_{i=1}^{k}\frac{T_{i.}^2}{n_i}-CF.$$

ويمكن الحصول علي SSE بطرح SSC من SSTO أى :

$$SSE = SSTO - SSC.$$

درجات الحرية سوف تصبح (N-1) لمجموع المربعات الكلية SSTO و (k-1) لمجموع مربعات متوسطات الأعمدة SSC و N-1-(k-1) = N-k لمجموع مربعات الخطأ.

مثال (١١-٢) أجريت تجربة لدراسة تأثير أربعة أنواع من الأدوية A, B, C, D علـــى الشفاء من مرض معين. البيانات معطاة في جدول (١١-٥) والتي تمثل عدد الأيام اللازمة للشفاء . استخدم طريقة تحليل التباين لاختبار ما إذا كان هناك فرق معنوي بين المتوسـطات عند مستوى معنوية $\alpha=0.05$.

جدول (١١-٥)

أنواع الأدوية			
A	B	C	D
3	7	3	10
4	8	2	12
3	4	1	8
5	10	2	5
	6	4	12
		2	10
		3	9
		1	

الحل . المطلوب اختبار فرض العدم :

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

ضد الفرض البديل :

واحد على الأقل من μ_i يختلف عن الباقي : H_1

$$\alpha=0.05.$$

$f_{.05}(3,20)=3.1$ والمستخرجة من جدول توزيـــع F في ملحـــق (٦) بدرجـــات حريـــة $\nu_1 = 3, \nu_2 = 20$. منطقة الرفض $F > 3.1$

$$SSTO = \sum_{i=1}^{k} \sum_{j=1}^{n} x_{ij}^2 - CF,$$

$$= 3^2 + 4^2 + \ldots + 10^2 + 9^2 - \frac{(134)^2}{24}$$

$$= 1030 - 748.17 = 281.83,$$

$$SSC = \sum_{i=1}^{k} \frac{T_{i.}^2}{n_i} - CF$$

$$= \frac{15^2}{4} + \frac{35^2}{5} + \frac{18^2}{8} + \frac{66^2}{7} - \frac{(134)^2}{24}$$

$$964.04 - 748.17 = 215.87,$$

$$SSE = 281.83 - 215.87 = 65.96.$$

جدول تحليل التباين معطي في جدول (١١-٦) .

جدول (١١-٦)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسط الأعمدة الخطأ	3 20	215.87 65.96	71.9567 3.298	21.818*
الكلي	23	281.83		

وحيث أن f المحسوبة (21.818) تقع في منطقة الرفض فإننا نرفض H_0 . أى أن هناك فرق معنوي بين المتوسطات .

(١١-٣) اختبار تجانس عدة تباينات :

Test for the Equality of Several Variances

ذكرنا في البند (١١-٢) أن هناك افتراضات أساسية وضرورية لإجراء تحليل التباين وهم : أن المجتمعات التي عددها k مستقلة وتتبع توزيعات طبيعية بمتوسطات $\mu_1, \mu_2, \ldots, \mu_k$ وتباين مشترك σ^2 . هناك العديد من الطرق المختلفة لاختبار فرض العدم :

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$$

ضد الفرض البديل :

H_1 : التباينات ليست كلها متساوية

اقترح Cochran [Winer et al (1991)] القيمة التالية :

$$c = \frac{\text{أكبر}\ s^2}{\Sigma s_i^2}$$

والتي تمثل قيمة للإحصاء C وذلك تحت فرض أن H_0 صحيح. القيم الحرجة $c_\alpha(v_1, v_2)$ للإحصاء C تستخرج من جدول Cochran في ملحق (٨) بدرجات حرية $v_1 = k, v_2 = n-1$ وذلك عند مستوى معنوية α=0.05 أو α=0.01 . منطقة الرفض $C > c_\alpha(v_1, v_2)$. إذا وقعت c في منطقة الرفض نرفض H_0 . بفرض

أن العينات التي عددها k ذات أحجام $n_1, n_2, \ldots, n_k$ (عدم تساوى حجوم العينـــات) وإذا كانت الأحجام متقاربة فيمكن استخدام أكبر n_i بدلاً من n في حساب درجات الحرية اللازمة لإيجاد $c_\alpha(v_1, v_2)$.

مثال (١١-٣) للبيانات في جدول (١١-٥) والخاصة بمثال (١١-٢) أختبر فـــرض العدم :

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$$

ضد الفرض البديل :

H_1 : التباينات ليست كلها متساوية

وذلك عند مستوى معنوية $\alpha=0.05$.

الحل . الجدول (١١-٧) يعطي تباين العينة لكل معالجة وعدد المشاهدات في كل معالجة .

جدول (١١- ٧)

المعالجة i	1	2	3	4
s_i^2	0.9167	5.0000	1.0714	5.9524
n_i	4	5	8	7

$$c = \frac{\text{أكبر } s^2}{\sum_{i=1}^{k} s_i^2} = \frac{5.9524}{12.9405}$$

$$= 0.459982,$$

وبما أن العينات التي عددها 4 ذات أحجام غير متساوية فسوف نأخذ $n = 8$ حيث 8 هـــي عدد المشاهدات في المعالجة رقم 3 (أكبر n_i) وعلى ذلك $v_1 = 4, v_2 = 8 - 1 = 7$ و $c_{.05}(4,7) = 0.5365$. منطقة الرفض $C > 0.5365$. وبما أن $c= 0.459982$ تقـع في منطقة القبول فإننا نقبل H_0 .

(١١-٤) اختبار دانكن للمدي المتعدد

نيومن

Duncan's Multiple Range Test

إذا كانت قيمة f المحسوبة من جدول تحليل التباين غير معنوية فــهذا يــدل علـــى أن الفروق بين متوسطات المعالجات ليست فروق حقيقية وإنما تعزى لمجرد الصدفة ، وبالتـــالي فإننا نقبل فرض العدم $H_0 : \mu_1 = \mu_2 = \ldots = \mu_k$. إذا كانت قيمة f معنوية فهذا يـــدل

على أن بعض الفروق بين متوسطات المعالجات أو كلها معنوية ، ولكن هـــذا الاختبـــار لا يوضح لنا أي من هذه الفروق معنوية ، ولذلك فإن الباحث لا بد أن يجري عدة مقارنـــات بين هذه المتوسطات وهذا ما يسمى بالمقارنات المتعددة. هناك عدة طـــرق تســـتخدم لهـــذا الغرض . سوف تقتصر دراستنا في هذا البند على اختبار دانكـــن للمقارنـــات المتعـــددة . يتلخص اختبار دانكن في إيجاد عدة فروق معنوية ذات قيم متزايدة والتي تتوقف حجمـــها على مدي البعد بين المتوسطات بعد ترتيبها وتتلخص خطوات تنفيذها على النحو التالي :

(أ) نرتب متوسطات المعالجات تنازلياً.

(ب) نوجد الخطأ المعياري للمتوسط $s_{\bar{x}} = \sqrt{\frac{MSE}{n}}$ حيث MSE هو متوسط مجموع مربعات الخطأ والذي يعتبر تقدير للتباين σ^2 ، ونحصل عليه من جدول تحليل التباين . وإذا كانت أحجام العينات للمعالجات غير متساوية فإن اختبار دانكن يسمح باستبدال n في صيغة $s_{\bar{x}}$ بالوسط التوافقي للقيم $n_1, n_2, \ldots, n_k$ حيث الوسط التوافقي :

$$\tilde{n} = \frac{k}{\frac{1}{n_1} + \frac{1}{n_2} + \ldots + \frac{1}{n_k}}$$

تحت شرط أن أحجام العينات تكون متقاربة من بعضها . هذا ويمكن استبدال n في صيغـــة $s_{\bar{x}}$ بالقيمة $n^{\cdot}$ حيث :

$$n^{\cdot} = \frac{2}{\frac{1}{n_{(1)}} + \frac{1}{n_{(k)}}}$$

و أن :

$n_{(1)}$ = حجم العينة المقابل لأصغر متوسط عينة .

$n_{(k)}$ = حجم العينة المقابل لأكبر متوسط عينة .

(ج) تستخرج قيم $q_\alpha(p, \nu)$ (تسمى أقل مدي معنوي قياســـي least significant studentized range) من جدول دانكن للمدى المعنوي في ملحق (٩) حيـــث $p = 2, 3, \ldots, k$ و α هي مستوى المعنوية و ν هي درجات حرية MSE.

(د) نحسب قيمة أقل مدى معنوي R_p least significant range وذلك بالنسبة لكل $p = 2,3, \ldots, k$ على النحو التالي :

$$R_p = q_\alpha(p,\nu)s_{\bar{x}}\ , p = 2,3,\ldots,k.$$

(هـ) نقارن الفروق بين متوسطات المعالجات ونبدأ بمقارنة الفرق بين أكبر متوسط وأقل متوسط بالقيمة R_k ثم نقارن الفرق بين أكبر متوسط وثاني أصغر متوسط بالقيمة R_{k-1} ونواصل هذه العملية وإلى أن تتم مقارنة كل الأزواج وعددها $\binom{k}{2} = k(k-1)/2$.

إذا كان الفرق المحسوب بين متوسطين يساوى أو أعلى من R_p فيكون ذلك الفرق معنويا. تلخص نتائج الاختبار بوضع خطوط مشتركة تحت المتوسطات التي لم تكن فروقها معنوية ، مع الإبقاء على ترتيب المتوسطات تنازليا:

لتوضيح طريقة دانكن للمدى المتعدد فسوف نستخدم البيانات الخاصة بمثال (١١-٢) ونتبع الخطوات التالية :

(أ) نرتب متوسطات المعالجات تنازلياً كالآتي :

$$\begin{matrix} \bar{x}_4 & \bar{x}_2 & \bar{x}_1 & \bar{x}_3 \\ 9.43 & 7.00 & 3.75 & 2.25 \end{matrix}$$

(ب) من جدول تحليل التباين (جدول (١١-٦)) فإن $MSE = 3.298$ بدرجات حرية $\nu = 20$. نوجد الخطأ المعياري للمتوسط $s_{\bar{x}} = \sqrt{\frac{MSE}{n}}$ وبما أن أحجام المعالجات غير متساوية فإننا نحسب الوسط التوافقي للقيم $n_1, n_2, \ldots, n_k$ كالآتي :

$$\tilde{n} = \frac{k}{\frac{1}{n_1}+\frac{1}{n_2}+\frac{1}{n_3}+\frac{1}{n_4}}$$

$$= \frac{4}{\frac{1}{4}+\frac{1}{5}+\frac{1}{8}+\frac{1}{7}} = \frac{4}{.7178571} = 5.5721,$$

$$MSE = 3.298, s_{\bar{x}} = \sqrt{\frac{MSE}{\tilde{n}}} = \sqrt{\frac{3.298}{5.5721}} = 0.7693.$$

يمكن تلخيص النتائج للحسابات السابقة في جدول (١١-٨) حيث قيم $q_{0.05}(p,20)$ تستخرج من جدول دانكن في ملحق (٩) حيث $\nu = 20$ ، $p = 2,3,4$.

جدول (١١-٨)

p	2	3	4
$q_{.05}(p,20)$	2.95	3.58	3.96
R_p	2.27	2.75	3.05

وبمقارنة قيم R_p بالفروق للمتوسطات المرتبة نحصل على الاستنتاجات الآتية

* وبما أن $\bar{x}_4 - \bar{x}_3 = 7.18 > R_4 = 3.05$ فإننا نستنتج أن الفرق بين $\bar{x}_4, \bar{x}_3$ معنوي .
* وبما أن $\bar{x}_4 - \bar{x}_1 = 5.68 > R_3 = 2.75$ فإننا نستنتج أن الفرق بين $\bar{x}_4, \bar{x}_1$ معنوي .
* وبما أن $\bar{x}_4 - \bar{x}_2 = 2.43 > R_2 = 2.27$ فإننا نستنتج أن الفرق بين $\bar{x}_4, \bar{x}_2$ معنوي .
* وبما أن $\bar{x}_2 - \bar{x}_3 = 4.75 > R_3 = 2.75$ فإننا نستنتج أن الفرق بين $\bar{x}_2, \bar{x}_3$ معنوي .
* وبما أن $\bar{x}_2 - \bar{x}_1 = 3.25 > R_2 = 2.27$ فإننا نستنتج أن الفرق بين $\bar{x}_2, \bar{x}_1$ معنوي .
* وبما أن $\bar{x}_1 - \bar{x}_3 = 1.5 < R_2 = 2.27$ فإننا نستنتج أن الفرق بين $\bar{x}_1, \bar{x}_3$ غير معنوي .

عادة تلخص الاستنتاجات بوضع خطوط تحت المتوسطات التي ليست بينها فروق معنوية وذلك كما يلي:

$$\bar{x}_4 \quad \bar{x}_2 \quad \underline{\bar{x}_1 \quad \bar{x}_3}$$

يتضح من الشكل السابق أن $\mu_4 > \mu_3, \mu_4 > \mu_1$ و $\mu_2 > \mu_1, \mu_2 > \mu_3, \mu_4 > \mu_2$.

يمكن تلخيص النتائج السابقة على النحو الموضح في جدول (١١-٩) حيث وضعت كل الفروق الممكنة بين المتوسطات داخل الجدول وتمت مقارنتها بقيم R_p المناسبة . يتضح من جدول (١١-٩) أن الفروق على كل قطر قيمة من أعلى اليسار إلى أدنى اليمين لها نفس قيم p . على سبيل المثال الفروق 2.43, 3.25, 1.5 تقع على قطر واحد ولها $p = 2$. القيمة الحرجة لهذه الفروق هي آخر قيمة في العمود الأخير (2.27) . أيضا الفروق , 5.68

4.75 تقع على قطر واحد ولها p=3 وتقارن بالقيمة 2.75 (القيمـــة الثانيـــة في العمـــود الأخير) . أخيرا الفرق 7.18 يقارن عند p=4 بالقيمة الحرجة 3.05 وهي القيمـة الأولي في العمود الأخير . النجمة * في الجدول تدل على أن الفرق معنوي وذلك عنـــــد اســـتخدام $\alpha=0.05$.

جدول (١١-٩)

المتوسطات	$\overline{x}_4 = 9.43$	$\overline{x}_2 = 7.00$	$\overline{x}_1 = 3.75$	$\overline{x}_3 = 2.25$	p	Rp
$\overline{x}_4 = 9.43$	---	2.43*	5.68*	7.18*	4	3.05
$\overline{x}_2 = 7.00$		---	3.25*	4.75*	3	2.75
$\overline{x}_1 = 3.75$			---	1.5	2	2.27
$\overline{x}_3 = 2.25$				---		

للسهولة يمكن تلخيص نتائج جدول (١١-٩) وذلك في جدول (١١-١٠) . نلاحظ أننا لم نرصد قيمة للفرق بين أي المتوسطين موضع المقارنة كما كنا نفعل من قبل بل رصدنا فقط نجمة .

جدول (١١-١٠)

	4	2	1	3
4		*	*	*
2			*	*
1				
3				

(١١-٥) التصنيف الثنائي ، مشاهدة واحدة في كل خلية

Two-Way Classification, Single Observation Per Cell

قد تصنف فئة من المشاهدات تبعا لصفتين معا. على سبيل المثال عندما يرغب الباحث في مجال الزراعة في دراسة تأثير الطرق المختلفة للزراعة (ثلاثة طرق) وكذلـــك الأوقـــات المختلفة للزراعة (مارس وفبراير ونوفمبر وأكتوبر) على إنتاجية محصـــول القصـــب. المشاهدات في هذه الحالة يمكن وضعها في جدول من ثلاثة صفوف وأربعة أعمدة حيث تمثل

الصفوف طرق الزراعة 3 ,2 ,1 وتمثل الأعمدة أوقات الزراعة (مارس وفبراير ونوفمبر وأكتوبر). يطلق على تقاطع أى صف مع أى عمود بالخلية وكل خلية تحتوي على مشاهدة واحدة. عموماً ، في حالة التصنيف الثنائي لمشاهدة واحدة يمكن وضع المشاهدات في جدول يتكون من r من الصفوف و c من الأعمدة كما هو موضح في جدول (١١-١١) حيث أن x_{ij} ترمز للمشاهدة في الصف رقم i والعمود رقم j . سوف نفترض أن x_{ij} قيم لمتغيرات عشوائية مستقلة لها توزيعات طبيعية بمتوسط μ_{ij} وتباين مشترك σ^2 . في جدول (١١-١١) $T_{i.}$ و $\bar{x}_{i.}$ ترمز للمجموع والمتوسط على التوالي لكل المشاهدات في الصف رقم i و $T_{.j}$ و $\bar{x}_{.j}$ ترمز للمجموع والمتوسط لكل المشاهدات في العمود رقم j و $T_{..}$, $\bar{x}_{..}$ ترمز للمجموع والمتوسط على التوالي لكل المشاهدات التي عددها rc . المتوسط لمتوسطات المجتمعات للصف رقم i ، $\mu_{i.}$ ، يعرف كالآتي :

$$\mu_{i.} = \frac{\sum_{j=1}^{c} \mu_{ij}}{c}.$$

جدول (١١-١١)

الصف	الأعمدة				المجموع	المتوسط
	1	2 ...	j ...	c		
1	$x_{11} \cdots$	$x_{12} \cdots$	$x_{1j} \cdots$	x_{1c}	$T_{1.}$	$\bar{x}_{1.}$
2	x_{21}	$x_{22} \cdots$	$x_{2j} \cdots$	x_{2c}	$T_{2.}$	$\bar{x}_{2.}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
i	$x_{i1} \cdots$	$x_{i2} \cdots$	$x_{ij} \cdots$	x_{ic}	$T_{i.}$	$\bar{x}_{i.}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
r	$x_{r1} \cdots$	$x_{r2} \cdots$	$x_{rj} \cdots$	x_{rc}	$T_{r.}$	$\bar{x}_{r.}$
المجموع	$T_{.1}$	$T_{.2} \cdots$	$T_{.j} \cdots$	$T_{.c}$	$T_{..}$	
المتوسط	$\bar{x}_{.1}$	$\bar{x}_{.2}$	$\bar{x}_{.j}$	$\bar{x}_{.c}$		$\bar{x}_{..}$

وبنفس الشكل ، المتوسط لمتوسطات المجتمعات للعمود رقم j و $\mu_{.j}$ يعرف كالآتي :

$$\mu_{.j} = \frac{\sum_{i=1}^{r} \mu_{ij}}{r}.$$

والمتوسط لمتوسطات المجتمعات التي عددها rc ، μ ، يعرف كالآتي :

$$\mu = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} \mu_{ij}}{rc}.$$

لتقدير ما إذا كان جزء من الاختلاف بين المشاهدات يرجع إلى الاختلاف بين الصفـــوف ، فإننا نختبر فرض العدم :

$$H_0' : \mu_{1.} = \mu_{2.} = \cdots = \mu_{r.} = \mu,$$

ضد الفرض البديل :

H_1': واحد على الأقل من $\mu_{i.}$ يختلف عن الباقي

وبنفس الشكل لتقدير ما إذا كان جزء من الاختلاف بين المشاهدات يرجع إلى الأعمـــدة ، فإننا نختبر فرض العدم :

$$H_0'' : \mu_{.1} = \mu_{.2} = \cdots = \mu_{.c} = \mu$$

ضد الفرض البديل :

H_1'': واحد على الأقل من $\mu_{.j}$ يختلف عن الباقي:

يمكن كتابته كل مشاهدة على الشكل :

$$x_{ij} = \mu_{ij} + \epsilon_{ij},$$

حيث ϵ_{ij} يقيس انحراف قيمة المشاهدة x_{ij} عن متوسط المجتمع μ_{ij}. الشـــكل المفصـل والشائع الاستخدام لهـــذه المعادلـــة (أو النمـــوذج) يمكـــن الحصـــول عليــه بوضـــع $\mu_{ij} = \mu + \alpha_i + \beta_j$ حيث α_i ترمز لتأثير الصف رقم i و β_j ترمز لتأثير العمود رقـم j . سوف نفترض أن تأثير الصفوف والأعمدة تجميعي additive (سوف نشـــرح ذلـــك بالتفصيل في البند التالي). وعلى ذلك يمكن إعادة كتابة x_{ij} على الشكل :

$$x_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij},$$

وذلك تحت القيود التالية :

$$\sum_{i=1}^{r} \alpha_i = 0 \ , \ \sum_{j=1}^{c} \beta_j = 0$$

وعلى ذلك :

$$\mu_{i.} = \frac{\sum_{j=1}^{c} (\mu + \alpha_i + \beta_j)}{c} = \mu + \alpha_i,$$

$$\mu_{.j} = \frac{\sum_{i=1}^{r} (\mu + \alpha_i + \beta_j)}{r} = \mu + \beta_j,$$

الآن اختبار فرض العدم :

$$H_0' : \mu_{1.} = \mu_{2.} = \cdots = \mu_{r.} = \mu,$$

ضد الفرض البديل : واحد على الأقل من $\mu_{i.}$ يختلف عن الباقي: H_1'

يكافئ اختبار فرض العدم :

$$H_0' : \alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$$

ضد الفرض البديل :

واحد على الأقل من α_i لا يساوي صفراً : H_1'

وبنفس الشكل اختبار فرض العدم :

$$H_0'' : \mu_{.1} = \mu_{.2} = \cdots = \mu_{.c} = \mu$$

ضد الفرض البديل :

واحد على الأقل من $\mu_{.j}$ يختلف عن الباقي : H_1''

يكافئ اختبار فرض العدم :

$$H_0'' : \beta_1 = \beta_2 = \ldots = \beta_c = 0$$

ضد الفرض البديل :

واحد على الأقل من β_j لا يساوي صفرا : H_1''

يعتمد الفرضين السابقين على المقارنة بين تقديرين مستقلين لمعلمة التباين المشترك σ^2 . هذين التقديرين نحصل عليهما بتجزئة مجموع المربعات الكلي للمشاهدات إلى ثلاثة مكونات كما يلي :

$$\sum_{i=1}^{r} \sum_{j=1}^{c} (x_{ij} - \overline{x}_{..})^2 = c \sum_{i=1}^{r} (\overline{x}_{i.} - \overline{x}_{..})^2 + r \sum_{j=1}^{c} (\overline{x}_{.j} - \overline{x}_{..})^2$$

$$+ \sum_{i=1}^{c} \sum_{j=1}^{c} (x_{ij} - \overline{x}_{i.} - \overline{x}_{.j} + \overline{x}_{..})^2$$

ويمكن التعبير عن الحدود في العلاقة السابقة باستخدام الرموز كالتالي :

SSTO = SSR + SSC + SSE ,

حيث مجموع المربعات الكلي هو :

$$SSTO = \sum_{i=1}^{r} \sum_{j=1}^{c} (x_{ij} - \bar{x}_{..})^2$$

ومجموع المربعات لمتوسطات الصفوف : sum of squares for rows means هو :

$$SSR = c \sum_{i=1}^{r} (\bar{x}_{i.} - \bar{x}_{..})^2,$$

ومجموع المربعات لمتوسطات الأعمدة هو :

$$SSC = r \sum_{j=1}^{c} (\bar{x}_{.j} - \bar{x}_{..})^2,$$

ومجموع مربعات الخطأ هو :

$$SSE = \sum_{i=1}^{r} \sum_{j=1}^{c} (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2.$$

التقدير الأول للمعلمة σ^2 ، يعتمد على $r-1$ درجات حرية ويعطى كالآتي :

$$MSR = \frac{SSR}{r-1}.$$

عندما $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$ فإن MSR يعتبر تقديرا غير متحيز للمعلمة σ^2 .

التقدير الثاني للمعلمة σ^2 يعتمد على c-1 درجات حرية ويعطي كالآتي :

$$MSC = \frac{SSC}{c-1}.$$

يعتبر التقدير MSC تقدير غير متحيز للمعلمة σ^2 إذا كان $\beta_1 = \beta_2 = \ldots = \beta_c = 0$.

التقدير الثالث للمعلمة σ^2 والذي يعتمد على $(r-1)(c-1)$ درجات حرية مستقل عن MSR و MSC ويعطي من الصيغة :

$$MSE = \frac{SSE}{(r-1)(c-1)},$$

وهو تقدير غير متحيز بصرف النظر عن صحة أو عدم صحة فرض العدم .

لاختبار فرض العدم H_0' فإننا نحسب النسبة :

$$f_1 = \frac{MSR}{MSE}$$

وهي قيمة لمتغير عشوائي F_1 يتبع توزيع F بدرجات حرية $(r-1)(c-1)$ و $(r-1)$ وذلك عندما يكون فرض العدم صحيح . سوف نرفض فرض العدم ، عند مستوى معنوية α ، عندما :

$$F_1 > f_\alpha(r-1,(r-1)(c-1)).$$

بنفس الشكل لاختبار فرض العدم H_0'' فإننا نحسب النسبة :

$$f_2 = \frac{MSC}{MSE}.$$

وهي قيمة لمتغير عشوائي F_2 يتبع توزيع F بدرجات حريــة $(r-1)(c-1)$ و $(c-1)$ وذلك عندما يكون فرض العدم صحيح. سوف نرفض فرض العدم ، عند مستوى معنوية α ، عندما :

$$F_2 > f_\alpha(c-1,(r-1)(c-1))$$

عملياً أولا نحسب SSTO و SSR و SSC ثم نحصل على SSE بطرح كــل مــن SSR و SSC من SST . أي أن SSE = SSTO – SSR – SSC . عادة درجــات الحرية المرتبطة بـــ SSE تحسب بطرح درجات الحرية الخاص بكل من SSR و SSC من درجات الحرية الخاصة بـــ SSTO. أي أن درجات الحرية الخاصة بـــ SSE هي:

$$(r-1)(c-1) = (rc-1)-(r-1)-(c-1).$$

عادة الصيغ المفضلة لحساب مجموع المربعات تعطي كالآتي :

$$SSTO = \sum_{i=1}^{r}\sum_{j=1}^{c} x_{ij}^2 - \quad CF,$$

$$SSR = \frac{\sum_{i=1}^{r} T_{i.}^2}{c} - \quad CF,$$

$$SSC = \frac{\sum_{j=1}^{c} T_{.j}^2}{r} - \quad CF,$$

حيث :

$$CF = \frac{T_{..}^2}{rc}.$$

جدول تحليل التباين في حالة التصنيف الثنائي بمشاهدة واحدة لكل خلية موضح في جـدول (١١-١٢) .

جدول (١١ – ١٢)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات	f المحسوبة
متوسطات الصفوف	$r-1$	SSR	$MSR=\frac{SSR}{r-1}$	$f_1=\frac{MSR}{MSE}$
متوسطات الأعمدة	$c-1$	SSC	$MSC=\frac{SSC}{c-1}$	$f_2=\frac{MSC}{MSE}$
الخطأ	$(r-1)(c-1)$	SSE	$MSE=\frac{SSE}{(r-1)(c-1)}$	
المجموع	$rc-1$	SST		

مثال (١١-٤) تعطي البيانات في جدول (١١-١٣) الدرجات التي حصل عليها ستة من الطلبة في ثلاثة مقررات والمطلوب :

(أ) هل هناك تفاوت في مقدرة الطلبة ؟

(ب) هل هناك تفاوت في صعوبة المقررات (استخدم مستوى معنوية $\alpha = 0.05$).

جدول (١١-١٣)

الطالب	المقرر		
	الرياضيات	اللغة الإنجليزية	اللغة الفرنسية
1	14	18	15
2	12	16	14
3	16	17	12
4	15	19	14
5	10	12	12
6	11	13	9

الحل .

(أ) $H'_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_6 = 0$

(ب) $H''_0 : \beta_1 = \beta_2 = \beta_3 = 0$

(أ) واحد على الأقل من α_i لا يساوي صفرا : H'_1

(ب) واحد على الأقل من β_j لا يساوى صفرا : H''_1

$f_{.05}(5,10)=3.33$ والمستخرجة من جـــدول توزيـــع F في ملحـــق (٦) بدرجـــات حريـــة $\nu_1 = 5$, $\nu_2 = 10$. أيضا $f_{.05}(2,10) = 4.1$. منطقة الرفض :

(أ) $F_1 > 3.33$

(ب) $F_2 > 4.1$

$$SSTO = \sum_{i=1}^{6} \sum_{j=1}^{3} x_{ij}^2 - CF$$

$$= 14^2 + 12^2 + \ldots + 12^2 + 9^2 - \frac{(249)^2}{18}$$

$$= 3571 - 3444.5 = 126.5,$$

$$SSR = \frac{\sum_{i=1}^{6} T_{i.}^2}{c} - CF$$

$$= \frac{47^2 + 42^2 + 45^2 + 48^2 + 34^2 + 33^2}{3} - \frac{(249)^2}{18}$$

$$= 3515.67 - 3444.5 = 71.17,$$

$$SSC = \frac{\sum_{j=1}^{3} T_{.j}^2}{r} - CF$$

$$= \frac{78^2 + 95^2 + 76^2}{6} - \frac{(249)^2}{18}$$

$$= 3480.83 - 3444.5 = 36.33.$$

جدول تحليل التباين موضح في جدول (١١-١٤) .

جدول (١١-١٤)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الصفوف	5	71.17	14.234	$f_1=7.492$
متوسطات الأعمدة	2	36.33	18.165	$f_2= 9.561$
الخطأ	10	19	1.9	
الكلـــي	17	126.5		

بما أن $f_1 = 7.492$ تقع في منطقة الرفض نرفض H_0' أي أن هناك تفاوت في مقدرة الطلبة.أيضا بما أن $f_2 = 9.561$ تقع في منطقة الرفض نرفض H_0'' أي أن هناك تفاوت في صعوبة المقررات .

(١١–٦) التصنيف الثنائي ، عدة مشاهدات لكل خلية

Two-Way Classification , Several Observations Per Cell

في البند (١١–٥) كان يفترض أن تأثير الصف والعمود تجميعي وهذا يكافئ أن $\mu_{ij} - \mu_{ij'} = \mu_{i'j} - \mu_{i'j'}$ أو $\mu_{ij} - \mu_{i'j} = \mu_{ij'} - \mu_{i'j'}$ لأي i, i', j, j' .

وهذا يعني أن الفرق بين متوسطات المجتمعات للعمودين j, j' متساوي لكل صف وأيضا الفرق بين متوسطات المجتمعات للصف i, i' متساوي لأي عمود. في كثير من التجارب لا يتحقق هذا الشرط وعلى ذلك فإن استخدام تحليل التباين الموضح في البند (١١–٥) يؤدى إلى استنتاجات خاطئة. للتوضيح وبالرجوع إلى التجربة الزراعية الخاصة بدراسة تأثير الأوقات المختلفة للزراعة (فبراير – مارس – أكتوبر – نوفمبر) وطرق الزراعة المختلفة 3 ,2 ,1 وبفرض أن نتيجة التجربة أوضحت أنه عند وقت الزراعة فبراير كان أعلي متوسط لمحصول القصب عند استخدام الطريقة 1 وأقل متوسط لمحصول القصب عند استخدام الطريقة 2 بينما عند وقت الزراعة مارس كان أعلي متوسط لمحصول القصب عند استخدام الطريقة 2 واقل متوسط لمحصول القصب عند استخدام الطريقة 1 . هذه الخاصية تعرف بالتفاعل بين أوقات الزراعة وطرق الزراعة وهي تكشف عما إذا كان لأوقات الزراعة آثار مختلفة على طرق الزراعة . كما لا يكون للتفاعل أثر إذا اتضح أن طرق الزراعة موضح البحث متناظرة لدي الأوقات المختلفة للزراعة .

لاختبار الفروق بين الصفوف والأعمدة عندما لا يتحقق الشرط التجميعي ، أي في وجود تفاعل **interaction** بين الصفوف والأعمدة ، لا بد من إيجاد تقدير غير متحيز للمعلمة σ^2 . افضل تقدير يمكن الحصول عليه إذا كررنا المشاهدات تحت نفس الظروف . للحصول على الصيغ العامة لتحليل التباين في هذه الحالة سوف نفترض الحالة التي تكون فيها عدد المشاهدات (المكررات) **replications** في كل خلية تساوي **n** . بفرض أن المشاهدات التي سوف نحصل عليها من التجربة ، في هذه الحالة ، يمكن ترتيبها في جدول ، مثل الجدول (١١–١٥) ، والذي يتكون من **r** من الصفوف و **c** من الأعمدة . وكل خلية تحتوي على **n** من المشاهدات . بفرض أن x_{ijk} ترمز للمشاهدة رقم **k** في الصف رقم **i** والعمود رقم **j** . المشاهدات التي عددها **rcn** موضحة في جدول (١١–١٥).

جدول (١١-١٥)

الصفوف	الأعمـــــــدة				المجموع	المتوسط
	1	2	. . .	c		
	x_{111}	x_{121}		x_{1c1}		
	x_{112}	x_{122}		x_{1c2}		
	.	.		.		
1	.	.	. . .	.	$T_{1..}$	$\bar{x}_{1..}$
	.	.		.		
	x_{11n}	x_{12n}		x_{1cn}		
	x_{211}	x_{221}		x_{2c1}		
	x_{212}	x_{222}		x_{2c2}		
	.	.		.		
2	.	.	. . .	.	$T_{2..}$	$\bar{x}_{2..}$
	.	.		.		
	x_{21n}	x_{22n}		x_{2cn}		
.	.	.		.		
.	.	.	. . .	.		
.	.	.		.		
	x_{r11}	x_{r21}		x_{rc1}		
	x_{r12}	x_{r22}		x_{rc2}		
	.	.		.		
r	.	.	. . .	.		
	.	.		.		
	x_{r1n}	x_{r2n}		x_{rcn}	$T_{r..}$	$\bar{x}_{r..}$
المجموع	$T_{.1.}$	$T_{.2.}$		$T_{.c.}$	$T_{...}$	$\bar{x}_{...}$
المتوسط	$\bar{x}_{.1.}$	$\bar{x}_{.2.}$		$\bar{x}_{.c.}$		

المشاهدات في الخلية رقم ij تمثل عينة عشوائية من الحجم n من مجتمع يفترض أنه يتبع توزيعـــاً طبيعياً بمتوسط μ_{ij} وتباين σ^2 . كل المجتمعات التي عددها rc يفترض أن لها تباين مشترك σ^2 . بقية الرموز المفيدة ، بعضها معطى في جدول (١١-١٥) ، يمكن توضيحها كالآتي :

$T_{ij.}$ = مجموع المشاهدات في الخلية رقم i,j

$T_{i..}$ = مجموع المشاهدات في الصف رقم i

$T_{.j.}$ = مجموع المشاهدات في العمود رقم j

$T_{...}$ = مجموع كل المشاهدات التي عددها rcn

$\bar{x}_{ij.}$ = متوسط كل المشاهدات في الخلية رقم ij

$\bar{x}_{i..}$ = متوسط كل المشاهدات في الصف رقم i

$\bar{x}_{.j.}$ = متوسط كل المشاهدات في العمود رقم j

$\bar{x}_{...}$ = متوسط كل المشاهدات التي عددها rcn

كل مشاهدة في جدول (١١-١٥) يمكن كتابتها على الشكل :

$$x_{ijk} = \mu_{ij} + \epsilon_{ijk},$$

حيث ϵ_{ij} يقيس انحراف قيمة المشاهدة x_{ijk} عن متوسط المجتمع . بفرض أن $(\alpha\beta)_{ij}$ يرمز لتأثير التفاعل للصف i مع العمود رقم j و α_i يرمز لتأثير الصف رقم i و β_j يرمز لتأثير العمود رقم j و μ تمثل المتوسط لكل المتوسطات فإن :

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

وعلى ذلك :

$$x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk},$$

تحت القيود التالية :

$$\sum_{i=1}^{r} \alpha_i = 0, \quad \sum_{j=1}^{c} \beta_j = 0, \quad \sum_{i=1}^{r} (\alpha\beta)_{ij} = 0, \quad \sum_{j=1}^{c} (\alpha\beta)_{ij} = 0$$

الفروض الثلاثة سوف نختبرها كالتالي :

(أ) $H_0' : = \alpha_1 = \alpha_2 = \ldots = \alpha_r = 0,$

H_1' : على الأقل واحد من α_i لا يساوي صفرا

(ب) $H_0'' := \beta_1 = \beta_2 = \ldots = \beta_c = 0,$

H_1'' : على الأقل واحد من β_j لا يساوي صفرا

(ج) $H_0''' : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \ldots = (\alpha\beta)_{rc} = 0,$

H_1''' : على الأقل واحد من $(\alpha\beta)_{ij}$ لا يساوي صفرا

كل اختبار من الاختبارات السابقة يعتمد على تقديرات مستقلة للمعلمة σ^2 وذلك بتجزئة مجموع المربعات الكلية للمشاهدات إلى أربعة مكونات كالآتي :

$$\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(x_{ijk}-\overline{x}_{...})=$$

$$cn\sum_{i=1}^{r}(\overline{x}_{i..}-\overline{x}_{...})^2+rn\sum_{j=1}^{c}(\overline{x}_{.j.}-\overline{x}_{...})^2$$

$$+n\sum_{i=1}^{r}\sum_{j=1}^{c}(\overline{x}_{ij.}-\overline{x}_{i..}-\overline{x}_{.j.}+\overline{x}_{...})^2$$

$$+\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(x_{ijk}-\overline{x}_{ij.})^2.$$

ويمكن التعبير عن مجموع المربعات في العلاقة السابقة باستخدام الرموز حيث مجموع المربعات الكلي هو :

$$SSTO=\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(x_{ijk}-\overline{x}_{...})^2,$$

ومجموع المربعات لمتوسطات الصفوف هو :

$$SSR=cn\sum_{i=1}^{r}(\overline{x}_{i..}-\overline{x}_{...})^2,$$

.ومجموع المربعات لمتوسطات الأعمدة هو :

$$SSC=rn\sum_{j=1}^{c}(\overline{x}_{.j.}-\overline{x}_{...})^2,$$

ومجموع المربعات للتفاعل بين الصفوف والأعمدة هو :

$$SS(RC)=n\sum_{i=1}^{r}\sum_{j=1}^{c}(\overline{x}_{ij.}-\overline{x}_{i..}-\overline{x}_{.j.}+\overline{x}_{...})^2,$$

ومجموع المربعات للخطأ هو :

$$SSE=\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n}(x_{ijk}-\overline{x}_{ij.})^2.$$

أي أن **SSTO = SSR + SSC + SS(RC) + SSE**

أيضا تجزأ درجات الحرية إلى :

$$rcn-1=(r-1)+(c-1)+(r-1)(c-1)+rc(n-1).$$

الآن يمكن الحصول على أربعة تقديرات غير متحيزة للمعلمة σ^2 عندما يكون H'_0, H''_0, H'''_0 صحيح وهم :

$$MSR=\frac{SSR}{r-1}\quad ,MSC=\frac{SSC}{c-1}\quad ,$$

$$MS(RC)=\frac{SS(RC)}{(r-1)(c-1)} \quad , \quad MSE=\frac{SSE}{rc(n-1)}.$$

لاختبار الفرض H_0' نحسب النسبة :

$$f_1=\frac{MSR}{MSE},$$

والتي تمثل قيمة للمتغير العشوائي F_1 والذي يتبع توزيع **F** بدرجات حرية $(r-1), rc(n-1)$عندما تكون H_0' صحيح. نرفض H_0' ، عند مستوى معنوية α ، عندما $F_1 > f_\alpha((r-1), rc(n-1))$. بنفس الشكل لاختبار الفرض H_0'' نحسب النسبة :

$$f_2=\frac{MSC}{MSE}$$

والتي تمثل قيمة للمتغير العشوائي F_2 والذي يتبع توزيع **F** بدرجات حرية $(c-1), rc(n-1)$ عندما تكون H_0'' صحيح. نرفض H_0'' ، عند مستوى معنوية α ، عندما $F_2 > f_\alpha((c-1), rc(n-1))$. أخيراً لاختبار الفرض H_0''' فإننا نحسب النسبة:

$$f_3=\frac{MS(RC)}{MSE}$$

والتي تمثل قيمة للمتغير العشوائي F_3 والذي يتبع توزيع **F** بدرجات حرية $(r-1)(c-1), rc(n-1)$ عندما تكون H_0''' صحيح . نرفض H_0'' ، عند مستوى معنوية α ، عندما $F_3 > f_\alpha((r-1)(c-1), rc(n-1))$ عادة يتم الحصول على مجموع المربعات من الصيغ التالية :

$$SSTO=\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n} x_{ijk}{}^2-CF,$$

حيث :

$$CF=\frac{T_{...}^2}{rcn} \text{ (معامل التصحيح) },$$

$$SSR=\frac{\sum_{i=1}^{r} T_{i..}{}^2}{cn}-CF,$$

$$SSC = \frac{\sum_{j=1}^{c} T_{.j.}^{2}}{rn} - CF,$$

$$SS(RC) = \frac{\sum_{i=1}^{r}\sum_{j=1}^{c} T_{ij.}^{2}}{n} - \frac{\sum_{i=1}^{r} T_{i..}^{2}}{cn} - \frac{\sum_{j=1}^{c} T_{.j.}^{2}}{rn} + CF$$

أما SSE فيمكن الحصول عليها من الصيغة التالية

SSE= SSTO – SSR – SSC – SS(RC).

الحسابات في مشكلة تحليل التباين ، في التصنيف الثنائية بعدة مشاهدات في كل خلية ، موضحة في جدول (١١ – ١٦).

جدول (١١–١٦)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الصفوف	r-1	SSR	$MSR = \frac{SSR}{r-1}$	$f_1 = \frac{MSR}{MSE}$
متوسطات الأعمدة	c-1	SSC	$MSC = \frac{SSC}{c-1}$	$f_2 = \frac{MSC}{MSE}$
التفاعل	(r-1)(c-1)	SS(RC)	$MS(RC) = \frac{SS(RC)}{(r-1)(c-1)}$	$f_3 = \frac{MS(RC)}{MSE}$
الخطأ	rc(n-1)	SSE	$MSE = \frac{SSE}{rc(n-1)}$	
الكلي	rcn-1	SSTO		

مثال (١١–٥) استخدمت ثلاثة مستويات من مبيد ما لمقاومة ثلاثة أجناس من حشرة Drosophila Pseudoobscura. تعطي البيانات في جدول (١١–١٧) معدلات الوفيات خلال فترة من الزمن . وتعتمد التجربة على خمسة مشاهدات في كل خلية.

المطلوب : (أ) اختبار معنوية الفروق بين مستويات المبيد.

(ب) اختبار معنوية الفروق بين الأجناس المختلفة من الحشرات .

(ج) التفاعل بين مستويات المبيد و الأجناس (مستوى معنوية α=0.05).

جدول (١١–١٧)

جدول (١١-١٧)

المستوي	الجنس		
	a_1	a_2	a_3
1	60,55,52,38,31	58,53,50,35,30	37, 43, 57, 60, 66
2	44,37,54,57,65	63,59,54,38,38	59,51,53,62,71
3	46,51,63,66,74	63,44,46,66,71	51,80,68,71,55

فروض العدم سوف تكون كالتالي :

(أ) $H_0' : \alpha_1 = \alpha_2 = \alpha_3 = 0,$

(ب) $H_0'' : \beta_1 = \beta_2 = \beta_3 = 0,$

(ج) $H_0''' : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \ldots = (\alpha\beta)_{33} = 0,$

الفروض البديلة سوف تكون كالتالي :

H_1' : على الأقل واحد من α_i لا يساوي صفراً

H_1'' : على الأقل واحد من β_j لا يساوي صفراً

H_1''' : على الأقل واحد من $(\alpha\beta)_{ij}$ لا يساوي صفرً

$3.23 \approx f_{.05}(2, 36)$ و $\alpha=0.05$ المستخرجة من جدول توزيع F في ملحق (٦). منطقة الرفض $F_1 > 3.23$.

$3.23 \simeq f_{.05}(2, 36)$ منطقة الرفض $F_2 > 3.23$

$2.61 \simeq f_{.05}(4, 36)$ منطقة الرفض $F_3 > 2.61$

البيانات في جدول (١١-١٧) يمكن تلخيصها في جدول (١١-١٨) . الآن :

$$SSTO = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{k=1}^{n} x_{ijk}{}^2 - CF$$

$$= 60^2 + 55^2 + \ldots + 71^2 + 55^2 - \frac{(2445)^2}{45}$$

$$= 139307 - 132845 = 6462,$$

$$SSR = \frac{\sum_{i=1}^{r} T_{i..}^{\ 2}}{cn} - CF$$

$$= \frac{725^2 + 805^2 + 915^2}{15} - \frac{(2445)^2}{45}$$

$$= 134058.33 - 132845 = 1213.33,$$

$$SSC = \frac{\sum_{j=1}^{c} T_{.j.}^{\ 2}}{rn} - CF$$

$$= \frac{793^2 + 768^2 + 884^2}{15} - \frac{(2445)^2}{45}$$

$$= 133341.93 - 132845 = 496.93,$$

$$SS(RC) = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} T_{ij.}^{\ 2}}{n} - \frac{\sum_{i=1}^{r} T_{i..}^2}{cn} - \frac{\sum_{j=1}^{c} T_{.j.}^2}{rn} + CF$$

$$= \frac{236^2 + 226^2 + \ldots + 325^2}{5} - 134058.33$$

$$-133341.93 + 132845 = 11.74,$$

$$SSE = SSTO - SSR - SSC - SS(RC)$$

$$= 6462 - 1213.33 - 496.93 - 11.74 = 4740.$$

جدول (١١-١٨)

المستوي	الجنس			
	a_1	a_2	a_3	المجموع
1	236	226	263	725
2	257	252	296	805
3	300	290	325	915
المجموع	793	768	884	2445

جدول تحليل التباين معطي في جدول (١١-١٩) .

جدول (١١-١٩)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الصفوف	2	1213.33	606.665	f_1=4.608
متوسطات الأعمدة	2	496.93	248.465	f_2= 1.887
التفاعـــل	4	11.74	2.935	f_3= 0.022
الخطـــأ	36	4740	131.666	
الكلـــي	44	6462		

من جدول تحليل التباين (جدول ١١-١٩) يمكن استنتاج :

(أ) نرفض H_0' لأن f_1 تقع في منطقة الرفض ، أي أن هناك فروق معنوية بين بمســـتويات المبيد.

(ب) نقبل H_0'' لأن f_2 تقع في منطقة القبول ، أي أنه لا يوجد فروق معنوية بـــين الأجناس.

(ج) نقبل H_0''' لأن f_3 تقع في منطقة القبول ، أي أنه لا يوجد تفاعل بين مســـتويات المبيد و أجناس الحشرة.

تماريـــــن :

-١- استخدم جدول تحليل التباين التالي ، عند مستوى معنويـــة α=0.05 ، في اختبـــار فرض العدم : $H_0 : \mu_1 = \mu_2 = \mu_3$.

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الأعمدة	2	70		
الخطـــأ	11	30		
الكلـــي				

-٢- في تجربة لدراسة المقاومة resistance لأربعة أنواع من الأسلاك اختيرت عينة عشـــوائية من 4 أسلاك من كل نوع وكانت MSC = 2573.3 و MSE = 1394.2. استخدم اختبار F عند مستوى معنوية $\alpha = 0.01$ لاختبار فرض العدم :

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

ضـــد الفرض البديل :

واحد على الأقل من $\mu_{i.}$ يختلف عن الباقي : H_1

-٣- يرغب طبيب متخصص في الأمراض النفسية في مقارنة ثلاث طرق لاختزال مستوى العدوانية في طلبة الجامعة. وقد تم إجراء اختبار HLT لقياس درجة العدوانية ، حيث تم اختيار الطلاب الذين لهم درجات عالية من العدوانية وعددهم 11 ، وقد تم اختيار 5 منهم للمعالجة بالطريقة A كما تم اختيار 3 للمعالجة B والثلاثة الباقين للمعالجة C . وقد استمرت المعالجة لمدة فصل دراسي . وفي نهاية الفصل الدراسي أعطي كل طالب اختبار HLT والنتائج في الجدول التالي :

A	74	63	68	80	79
B	54	74	71		
C	79	95	57		

(أ) أوجد جدول تحليل التباين ؟

(ب) اختبر معنوية الفروق بين المعالجات الثلاثة ؟

-٤- أجريت تجربة لدراسة تأثير أربعة طرق على زمن الطيران لبعوضة الملاريا (زمن الطيران خلال 24 ساعة)؟ الطرق هي IRN و IRC و IRS و C (معالجة المراقبة). النتائج التي تم الحصول عليها هي :

$$\bar{x}_{1.} = 4.39(IRS)\ ,\ \bar{x}_{2.} = 4.52(IRC),$$

$$\bar{x}_{3.} = 5.49(IRN)\ ,\ \bar{x}_{4.} = 6.3(C),$$

$$\sum_{i=1}^{4}\sum_{j=1}^{10} x_{ij}^2 = 1911.91.$$

استخدم جدول تحليل التباين لاختبار معنوية الفروق بين المعالجات المختلفة عند مستوى معنوية $\alpha=0.05$.

-٥- أكمل جدول تحليل التباين التالي وأختبر معنوية الفروق بين الأنواع عند مستوي معنوية $\alpha=0.01$.

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
الأنواع (متوسطات الأعمدة)	3			
الخطـــا			14.7137	

الكلــي	19	310.50076		

٦- تعطي البيانات التالية أربعة قراءات لكل نوع من الطائرات حيث تمثل كل قــراءة ، الزمن بالساعة ، الذي قطعته الطائرة من المدينة أ إلى المدينة ب :

نوع الطائرة	1	5.0	5.5	5.9	5.7
	2	7.0	8.0	8.1	8.2
	3	4.9	4.8	4.5	4.6
	4	7.1	7.4	7.6	7.8

(أ) أكتب جدول تحليل التباين ؟

(ب) أختبر معنوية الفروق بين الأنواع الأربعة ؟

(ج) في حالة إذا كانت الفروق معنوية أى من هذه الأنواع يختلف عن الآخر ؟

٧-البيانات التالية تمثل كمية Fe لأربعة أنواع من سبائك الحديد :

(1 Carbonate – 2 Silicate – 3 Magnetite , 4 Hematite)

أنواع السبائك	1	20.5	28.1	27.8	37.0	28.0		
		25.2	15.5	27.1	20.5	31.3		
	2	26.3	24	26.2	20.2	23.7		
		34.5	17.1	26.8	23.7	24.9		
	3	29.5	34	27.5	29.4	24.9		
		26.2	29.9	29.5	30.0	35.5		
	4	36.5	44.2	34.1	30.3	31.4	33.1	34.1
		32.9	36.3	25.5				

(أ) أختبر الفرضية $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ عند مستوى معنوية $\alpha=0.05$.

٨- لتقدير فيما إذا كان هناك فروق معنوية في كمية النتروجين في ثلاثة مواقع مختلفة مــن بحيرة ، أخذت 8 عينات من كل موقع . البيانات في الجدول التالي مقاسة milligrams / 100 grams .

الموقـــع		
A	B	C
222	326	263
300	275	360
262	218	221
264	207	198
200	272	211
211	268	266
267	308	312
326	229	299

المطلوب تحليل البيانات للفروق المعنوية وإذا كانت قيمة f معنوية اختبر الفروق المعنوية بين كل أزواج متوسطات المواقع.

-٩- يعيش طائر معين في ثلاثة مناطق جغرافية ، وقد اختيرت عينة عشوائية من كل منطقة في المناطق الثلاثة وتم قياس طول المنقار بالملليمتر لاقرب رقم عشري والبيانات في الجـــدول التالي :

الموقع		
A	B	C
4.2	3.8	3.0
3.3	4.2	3.4
2.8	5.0	4.4
4.2	4.5	4.5
3.7	5.2	
4.4		
3.5		

المطلوب تحليل البيانات للمعنوية الإحصائية بين المناطق المختلفة وإذا كـــانت f المحســوبة معنوية أختبر الفروق المعنوية الإحصائية بين كل أزواج متوسطات المواقع .

-١٠- يرغب باحث في العلوم البيولوجية في دراسة تأثير المستويات المختلفة من الإثيــانول على زمن النوم. اختيرت عينة عشوائية من 5 فأر (متساوية في الوزن والعمر) لكل معالجة . وقد تم حقن كل فأر . وقد تم تسجيل سرعة حركة العينة في زمن النوم rapid eye movement sleep time خلال فترة 24 ساعة والبيانات كما يلى :

0 g/kg	88.6	73.2	91.4	68.0	75.2
1 g/kg	63.0	53.9	69.2	50.1	71.5
2 g/kg	44.9	59.5	40.2	56.3	38.7
4 g/kg	31.0	39.6	45.3	25.2	22.7

(ب) أوجد جدول تحليل التباين؟

(ت) أختبر معنوية الفروق بين المعالجات الثلاثة ؟

(ج) أستخدم اختبار Cochran للتأكد من تجانس التباين عند مستوى معنوية $\alpha=0.01$

-١١- أجريت مقارنة لخمس طرق لتدريس مقرر الإحصاء الوصفي. الطرق الخمسة هي R تنفيذ برامج على الورق و L | R تنفيذ برامج على الورق + محاضرات C, تطبيق على الحاسب الآلي و L | C التطبيق مع المحاضرات و L | D تدريب ومناقشات. وقد تم اختبار عينة عشوائية من 9 طلاب لكل طريقة وبعد نهاية المقرر أعطي امتحان لمدة ساعة لكل الطلبة والنتائج في الجدول التالي :

	$\bar{x}_{i.}$	s_i
L | D	29.3	4.99
R	28.0	5.33
L | R	30.2	3.33
C	32.4	2.94
L | C	34.2	2.74

هل تدل هذه البيانات على أن هناك فروق معنوية بين متوسطات الطرق الخمسة في التدريس؟ (مستوى المعنوية $\alpha=0.05$).

-١٢- لمقارنة أسعار أحجام مختلفة من الثلاجات في مدينة ما تم الحصول على النتائج التالية :

النوع	حجم العينة	$\bar{x}_{i.}$	s_i^2
Frigidaire	18	123.21	18.71
GE	28	418.13	21.15
Whirlpool	19	421.72	17.8

أختبر معنوية الفروق بين متوسطات الأنواع المختلفة من الثلاجات عند مستوى معنوية $\alpha=0.05$.

-١٣- الجدول التالي يعطى عدد الأميال في الجالون لثلاثة أنواع من البنزين استخدمت لمدة 5 أيام . أختبر فرض العدم أن متوسطات البنزين متساوية عند مستوى معنوية $\alpha=0.05$.

البنزين						
	1	11	13	14	15	12
	2	16	15	14	15	15
	3	18	16	15	16	17

١٤- بفرض أن خمسة أنواع من الخلطات الغذائية diets ،والمحتوية على مصادر مختلفة من الكربوهيدرات ، غذيت بها خمسة مجموعات من الفئران. البيانات التالية تمثل كمية DNA في كبد كل فأر (مقاسه بالمليجرام لكل جرام من وزن الكبد).

	$\bar{x}_{i.}$
نشـــا	2.58
سكروز	2.63
فركتوز	2.13
جلوكوز	2.03
مالتوز	2.49

حيث $\sum_{i=1}^{5}\sum_{j=1}^{4} x_{ij}^2 = 183.4$. هل تدل النتائج السابقة على أن هناك فروق معنوية بـين المتوسطات (عند مستوى معنوية $\alpha=0.05$).

١٥- في تجربة لمقارنة ثلاثة طرق لتدريس مقرر في الرياضيات ، الطريقة الأولى كان التدريس فيها نظري – الطريقة الثانية كان التدريس فيها باستخدام شرائط الفيديو – الطريقة الثالثة كان التدريس فيها بالتطبيق على الحاسب الآلي. وقد تم اختبار عينة عشوائية من 5 طلاب لكل طريقة وطبقت كل طريقة لمدة فصل دراسي. وفي نهاية الفصل الدراسي أعطي كل طالب نفس الامتحان. الدرجات في الجدول التالي :

الطريقة	1	86	82	94	77	86
	2	90	79	88	87	96
	3	78	70	65	74	63

أختبر معنوية الفروق بين متوسطات الطرق الثلاثة عند مستوى معنوية $\alpha=0.05$.

١٦- البيانات في الجدول التالي تمثل الدرجات النهائية في مادة الإحصاء والتي حصل عليها ثلاثة مجموعات الكلية من نفس الفرقة ثم تدريسهم على يد ثلاثة محاضرين A,B,C والمطلوب التحقق من معنوية الفروق بين هذه المجموعات عند مستوى معنوية $\alpha=0.05$.

	المجاميع	
A	B	C
72	88	97
88	78	68
82	49	91
42	40	70
81	50	70
72	86	41
65	74	59

-١٧- الجدول التالي يعطى الإنتاج اليومي لثلاثة آلات ، وقد تم تسجيل الإنتاج اليومي لكل آلة في فترة أربعة أيام اختيرت عشوائيا . المطلوب اختبار معنوية الفروق بين متوسطات الآلات عند مستوى معنوية $\alpha=0.05$.

الآلـــة		
A	B	C
110	115	120
108	114	121
107	116	122
106	117	123

-١٨- بفرض أن أربعة أنواع من الفيتامينات A, B, C, D وتغذي عليها أربعة مجموعات من الأطفال متشابهين تماما (أربعة عينات عشوائية) وكانت الزيادة في وزن كل مجموعة كالآتي :

الفيتامينات			
A	B	C	D
2	2	4	4
3	1	5	3
3	2	4	2
	3	4	3

(أ) قدر متوسطات الزيادة في الوزن عند الفيتامينات الأربعة ؟

(ب) أوجد التباين لكل مجموعة ، وأجري اختبار **Cochran** .

(ج) استخدم طريقة تحليل التباين لاختبار ما إذا كان هناك فروق معنوية بين المتوسطات عند مستوى معنوية $\alpha=0.01$.

-١٩- صممت أربعة مقاعد خاصة بالسيارات وبها حزام الأمان وذلك لإعطاء أحسن حماية من حوادث الرأس عند السرعة **35 mph** أو أقل وقد أجري اختبار لمحاكاة الحوادث وتم الحصول على الدرجات التالية :

المقاعد			
A	B	C	D
37	49	32	40
41	37	33	48
44	40	41	40
48	38	37	41
50	51	48	37
44	42	37	40

أوجد جدول تحليل التباين واختبر معنوية الفروق بين متوسطات المقاعد عند مستوى معنوية $\alpha=0.05$.

٢٠- لتقدير إمكانية تأثير العامل الوراثي على ضغط الدم تم تحديد أربعة أنواع من الفئران A, B, C, D وقد اختيرت عينة عشوائية من كل نوع وتم قياس ضغط الدم لكل فأر وبيانات التجربة في الجدول التالي :

الأنواع			
A	B	C	D
83	85	88	89
81	84	94	85
85	84	91	88
59	91	92	92
85	88	95	84
92	89	88	85
91	91	87	40
80	89	91	93
79	86	90	90
82	87	93	89

المطلوب تحليل البيانات للفروق الإحصائية بين المتوسطات. وإذا كانت قيمة f معنوية أختبر للفروق الإحصائية بين كل أزواج المتوسطات (عند مستوى معنوية $\alpha=0.01$).

٢١- البيانات التالية تمثل النسبة المئوية (في المتوسط) لكحل المثيل methyl alcohol (لكل زجاجة) والتي تم تحليلها من قبل أربعة معامل .

المعمل	1	85.06	82.25	84.87
	2	48.99	89.28	84.88
	3	89.48	84.72	85.1
	4	84.1	84.55	84.05

أ- أكتب جدول تحليل التباين؟

ب- هل توجد فروق معنوية بين المعامل (عند مستوي معنوية $\alpha=0.01$).

-٢٢- جربت خمسة أنواع من الأسمدة على عدد من القطع المزروعة قمح . وقد طبقت المعالجة (١) على أربعة قطع والمعالجة (2) و (3) على 6 قطع والمعالجة (4) على 8 قطع. أما المعالجة (5) فقد طبقت على 3 قطع والمحصول لكل ياردة مربعة في الجدول التالي :

الأنواع				
(1)	(2)	(3)	(4)	(5)
78.9	63.5	79.1	87.0	75.9
72.3	74.1	90.3	91.0	77.2
81.1	75.5	85.6	91.2	81.2
85.7	80.8	81.4	75.3	
	71.3	74.5	79.4	
	79.4	95.3	80.7	
			82.8	
			89.6	

هل يوجد فروق بين تأثير الأنواع المختلفة من الأسمدة ؟ استعمل مستوى معنوية $\alpha=0.05$.

-٢٣- للمقارنة بين أربعة أنواع من مشروب بارد (مصنعة تبعاً لمكسب اللون المضاف (بدون لون – أحمر – برتقالي – أخضر) وقد تم توزيع كل نوع عشوائياً على خمسة مواقع وسجل عدد حالات البيع لكل 1000 شخص في الموقع خلال فترة الدراسة والبيانات في الجدول التالي :

أنواع المشروب			
بدون لون	أحمر	برتقالي	أخضر
26.5	31.2	27.9	30.8
28.2	28.3	25.1	29.6
25.1	30.8	28.5	32.4
29.1	27.9	24.2	31.7
27.2	29.6	26.5	32.8

(أ) قدر متوسطات المبيعات عند الألوان المختلفة ؟

(ب) أوجد جدول تحليل التباين وأجري اختبار **Cochran** للتجانس ؟

(ج) استخدم طريقة تحليل التباين لاختبار ما إذا كان هناك فروق معنوية بين المتوسطات عند مستوى معنوية $\alpha=0.01$.

-٢٤- قام مسئول الإنتاج في مصنع للبطاريات باختبار عمر البطارية بالساعة لأربعة أنواع من البطاريات التي تستخدم في تشغيل أجهزة الراديو المحمول . وقد تم استخدام 6 أجهزة

لكل نوع من البطاريات وتشغيل الأجهزة على موجة عالية لفترة معينة . وتم الحصول على البيانات التالية :

أنواع البطاريات			
(1)	(2)	(3)	(4)
5/5	4.7	6.1	5.5
5/0	3.9	5.7	5.1
5/2	4.3	5.0	4.3
5.3	4.5	5.3	4.1
4.8	4.3	6.3	5.1
5.0	4.0	5.8	4.2

هل هناك تفاوت في متوسط أعمار البطاريات للأنواع ؟

٢٥- للمقارنة بين أربعة أنواع من الحبوب من حيث كمية الثيامين (mg/g) قام باحث باختبار 6 عينات من كل نوع وتم الحصول على البيانات التالية :

Wheat	5.2	4.2	6.0	6.1	6.7	5.8
Barley	6.5	8.0	6.1	7.5	5.9	5.6
Maize	5.8	4.7	6.4	4.9	6.0	5.2
Watt	8.3	6.1	7.8	7.0	5.5	7.2

فهل تدل هذه البيانات على الاختلاف في متوسط الثيامين بين الأنواع الأربعة من الحبوب ؟

٢٦- أجريت تجربة لدراسة التأثير السام لثلاثة أنواع من الكيمائيات A, B, C على جلد الفئران وذلك عن طريق معالجة بوصة مربعة من الجلد بالمادة الكيمائية وإعطاء درجات من 5 إلى 15 على حسب درجة التأثير على الجلد وقد تم اختبار عينة من 7 أنسجة لكل معالجة و البيانات في الجدول التالي :

A	9	6	5	7	5	6	6
B	9	9	8	8	7	7	7
C	6	5	6	8	5	5	7

(أ) أكتب جدول تحليل التباين ؟

(ب) أختبر معنوية الفروق لمتوسطات المعالجات ؟

٢٧- قام باحث في مجال علوم الأغذية بدراسة تأثير الكميات المختلفة من اللبن المضاف إلى عجينه الكيك (منخفض - متوسط - عالي) على حجم الكيك المخبوزة (مقاس milliliters per 100 grams) و البيانات في الجدول التالي :

	1	351	369	381	380	370	358
الكيك	2	390	394	406	407	415	375
	3	398	409	415	399		

(أ)أوجد التباين لكل مجموعة وأجرى اختبار **Cochran** للتجانس ؟

(ب)استخدم طريقة تحليل التباين لاختبار ما إذا كان هناك فروق معنوية بين المتوسطات عند مستوى معنوية $\alpha=0.01$.

-٢٨- البيانات التالية تم الحصول عليها من تجربة لمقارنة ثلاثة أنواع من المبيدات استخدمت في رش أماكن مصابة بحشرة الخنفساء . كل مشاهدة تمثل عدد الوفيات من الخنافس في منطقة محددة تحتوى على هذا المبيد .

1 11, 9, 13, 11
2 6, 28, 31, 27, 30,33
3 19, 23, 19, 21, 20

حلل البيانات للفروق المعنوية في معدل الوفاة بين الأنواع المختلفة من المبيد وذلك عند مستوى معنوية $\alpha=0.05$.

-٢٩- لاختبار فاعلية خمسة أنواع من الأسمدة على إنتاج الذرة الصفراء وتم الحصول على البيانات التالية (نتائج المحصول مقاس بالكيلوجرامات لكل قطعه).

السماد				
(1)	(2)	(3)	(4)	(5)
40	38	44	41	34
45	40	42	43	35
46	38	40	40	34
49	44	34	40	23

المطلوب التحقق من مواقع الفروق بين متوسطات المجموعات المختلفة عند مستوى معنوية $\alpha=0.05$.

-٣٠- لدراسة تأثير المستوى الاجتماعي لطلاب إحدى الجامعات على مستواهم العلمي ، قام باحث بتحديد ثلاثة مستويات اجتماعية واختيار من كل مستوى عينة عشوائية من الطلاب وسجل المعدل التراكمي **GRA** لكل منهم . البيانات في الجدول التالي :

المستوي		
المستوى الأول	المستوى الثاني	المستوى الثالث
2.15	3.22	2.22
2.71	2.55	3.44
1.71	3.97	3.99
2.16	3.88	3.52
3.13	3.87	3.66

(أ)أختبر معنوية الفروق بين المستويات الاجتماعية الثلاثة وذلك عند مستوى معنوية $\alpha=0.05$.

(ب) في حالة ما إذا كانت الفروق معنوية ، أي هذه المستويات يختلف عن الأخر .

٣١- لمقارنة أربعة أنواع مختلفة من الصناديق بالنسبة لقوة الضغط Compression strength (مقاس بالرطل) تم الحصول على البيانات التالية :

الصندوق	1	655.5	788.3	734.3	721.4	679.1	699.4
	2	782.2	772.5	786.9	686.1	732.1	774.8
	3	737.1	639.0	696.3	671.7	717.2	721.7
	4	535.1	628.7	542.4	559	586.9	52.0

أختبر معنوية الفروق بين الصناديق الأربعة عند مستوى معنوية $\alpha=0.01$.

٣٢-الجدول التالي يمثل المبيعات لثلاثة أنواع من الفطائر (المبيعات بالدولار) تم عرضها في 16 مركزاً للبيع خلال أسبوعين

الفطائر	1	2161	1769	2548	1782	
	2	2379 1689	1419	1119	1208	1962
	3	1479 2215	1024	1598	4613	1913

(أ) المطلوب تحليل هذه البيانات بأسلوب تحليل التباين عند مستوى معنوية $\alpha=0.01$.

٣٣- لمقارنة أسعار رغيف الخبز (من نوع ما) في أربعة مواقع في مدينة ما ، اختيرت عينة عشوائية من أربعة مخابز في كل منطقة وتم تسجيل المبيعات في كل مخبز لفترة معينة والبيانات في الجدول التالي :

الموقع	1	55	63	65	61
	2	58	61	55	58
	3	54	51	58	58
	4	69	70	71	77

(أ) أكتب جدول تحليل التباين ؟

(ب) أختبر معنوية الفروق بين متوسطات المواقع ؟

(ج) في حالة ما إذا كانت الفروق معنوية أي من هذه المواقع يختلف عن الآخر ؟

٣٤- أخذت عينات من الماء من مواقع مختلفة من نهر لتقدير فيما إذا كانت متساوية في كمية الأوكسجين المذاب والذي يعتبر مقياس لتلوث المياه ، وقد أخذت عينة عشوائية من كل موقع والبيانات في الجدول التالي :

	1	5.9	6.1	6.3	6.1	6.0
	2	6.3	6.6	6.4	6.4	6.5
الموقع	3	4.8	4.3	5.0	4.7	5.1
	4	6.0	6.2	6.1	8.5	

(أ) هل تدل هذه البيانات على أن هناك اختلاف في متوسط ذوبان الأوكسجين بين المواقع الأربعة (عند مستوى معنوية $\alpha=0.05$).

(ت) استعمل اختبار دنكن للمقارنة بين المتوسطات للمواقع المختلفة ؟

-٣٥- قام باحث في مجال الزراعة بدراسة لمقارنة معدل النمو لنبات مائي في أربعة مواقع. جزء من الدراسة تناول طول الورقة لهذا النبات اختيرت عينات عشوائية من كل موقع. تعطى البيانات التالية متوسط طول الورقة لكل نبات (مقاسه بالسنتيمتر) لعينة عشوائية من 10 ورقات لكل نبات .

	1	5.7	6.3	6.1	6.0	5.8	6.2
	2	6.2	5.3	5.7	6.0	5.2	5.5
الموقع	3	5.4	5.0	6.0	5.6	4.9	5.2
	4	3.7	3.2	3.9	4	3.5	3.6

هل توجد فروق معنوية بين المواقع عند مستوى معنوية $\alpha=0.05$ ؟

-٣٦- بفرض أن خمسة أنواع من الخلطات diets غذيت بها خمسة مجموعات من الفئران متشابهين تماماً وسجلت الزيادة في الوزن والنتائج في الجدول التالي

الخلطة	n_i	$\bar{x}_{i.}$	s_i^2
Red , Maple '74	13	1.134	0.0252
Red , Oak / red maple	10	1.148	0.253
Red Maple '75	20	1.159	0.0179
Red Oak	16	1.191	0.0200
Red Oak/white pine	16	1.217	0.016

أ- أحسب $\bar{x}_{..}$.

ب-المطلوب تحليل هذه البيانات بأسلوب تحليل التباين؟

-٣٧- قام باحث بعمل مقارنة لستة أنواع من الزبد الصناعي تخص الأحماض الدهنية الغير مشبعة والبيانات في الجدول التالي:

Imperial	14.1	13.6	14.4	14.3	
Parkay	12.8	12.5	13.4	13	12.3
Blue Bounet	13.5	12.7	12.6	13.9	
Chiffron	16.8	17.2	16.4	17.3	
Mazola	16.8	17.2	16.4	17.3	18.0
Fleischmann's	18.1	17.2	18.7	18.4	

المطلوب تحليل هذه البيانات بأسلوب تحليل التباين ؟

-٣٨- أجرى باحث في مجال العلوم الميكروبيولجي اختبار الظـاهرة التفـاعل الضوئـي photoreactive phenomenon في البكتريا عند تعريضها للأشعة : (1) العينـة الأولي تم تعريضها للضوء المرئي لمدة 10 دقائق قبل وضعها في الحضانة (2) العينة الثانية تم تعريضها للضوء المرئي لمدة 5 دقائق قبل وضعها في الحضانة (3) العينة الثالثة وضعـت في الحضانـة بدون تعريضها لضوء البيانات التالية تمثل عدد مستعمرات البكتريا وذلك لخمسة تكرارات لكل معالجة :

(1) 90, 100, 75, 70, 64
(2) 72, 81, 53, 48,55
(3) 45, 40, 10, 23, 32

المطلوب :

(أ) اكتب جدول تحليل التباين

(ب) اختبر معنوية الفروق بين المعالجات المختلفة

(ث) في حالة ما إذا كانت الفروق معنوية أى من هذه المعالجات يختلف عن الآخر؟

-٣٩- بالرغم من أن الشاي يعتبر من أكثر المشروبات انتشارا بعد الماء إلا أن الكثـير لا يعلم عن قيمته الغذائية nutritional value . فهناك folocin وهو واحد من مجموعـة فيتامين B والذي يوجد بكميات معنوية في الشاي . استخدمت طرق تحليل دقيقة لحسـاب كمية folocin في الشاي في عينات عشوائية اختبرت من أربعة أنواع من الشاي الأخضـر . البيانات في الجدول التالي :

نـوع الشاي							
1	7.9	6.2	6.6	8.6	8.9	10.1	9.6
2	5.7	7.5	9.8	6.1	8.4		
3	6.8	7.5	5	5	5.3	6.1	
4	4.6	7.1	7.9	7.9	5.4		

هل البيانات السابقة تدل على أن متوسط folocin واحد في كل أنواع الشاي ؟ (وذلـك عند مستوى معنوية $\alpha=0.01$).

-٤٠- تعطى البيانات التالية محصول الطماطم (kg/plot) أثناء الزراعة بالمعروض لأربعـة مستويات من التوصيل الكهربائي electrical conductivity .

المستويات					
1.6	59.5	53.3	56.8	63.1	
3.8	55.2	59.1	52.8	54.9	
6.0	52.7	48.8	53.9	49	
10.2	44.6	48.5	41	47.3	46.1

والمطلوب اختبار معنوية الفروق بين متوسطات المستويات عند مستوى معنوية $\alpha=0.01$.

-٤١- يعتقد باحث أن درجة الحرارة والملوحة عاملان مهمان في إنتاجية محصول الجمبري . لذلك تم تصميم تجربة ذات عاملين ، العامل الأول (الملوحة) له ثلاثة مستويات والعامل الثاني درجة الحرارة وله ثلاثة مستويات. وقد تم تربيه الجمبري في المستويات السابقة من الملوحة والحرارة . وتسجيل محصول الجمبري في الجدول التالي (البيانات مقاسه بعدد التانكات الكبيرة من سعة 80 جالون).

الحرارة	الملوحة		
	A	B	C
60°F	3	5	4
70°F	11	10	12
80°F	16	2	17

-أ- أوجد جدول تحليل التباين لهذه التجربة ؟

-ب- هل تبين البيانات أن هناك فرقاً معنوياً بين متوسطات المعالجات عند مستوى معنوية $\alpha=0.01$.

-٤٢- تعطى البيانات في الجدول التالي مبيعات منظف لغلافين وثلاثة تركيبات مختلفة والمطلوب الإجابة على التساؤلات الآتية :

(أ) هل يختلف متوسط المبيعات باختلاف تركيبة الصابون؟

(ب) هل يختلف متوسط المبيعات باختلاف نوع الغلاف ؟

الغلاف	التركيبة		
	تركيبة (1)	تركيبة (2)	تركيبة (3)
غلاف 1	83	75	79
غلافه 2	74	75	78

-٤٣- في دراسة على تأثير عادم السيارات على تلوث الهواء أخذت عينات من الهواء عند أزمنة مختلفة وعند مواقع مختلفة وتم تحليلها لمعرفة كمية المادة المسببة للتلوث الموجود في الهواء مقاسه mg/m^3 . البيانات في الجدول التالي

الوقت	الأزمنة				
	1	2	3	9	5
أكتوبر ١٩٧٢	76	67	81	56	51
يناير ١٩٧٦	82	69	96	59	70
مايو ١٩٧٦	68	59	67	54	42
سبتمبر ١٩٧٦	63	56	64	58	37

(أ) أختبر معنوية الفروق بين المواقع ؟

(ب) أختبر معنوية الفروق بين الأوقات ؟

٤٤- في تجربة لمقارنة أربعة أنواع من الطلاء وثلاثة أنواع من البكرات (المستخدمة في الطلاء) استخدم جالون من كل نوع وأستخدم بكرة للطلاء وتعطى البيانات التالية عدد الأقدام المربعة التي غطاها الطلاء .

نوع الدهان	البكرة			
	1	454	446	451
	2	446	444	442
	3	439	442	444
	4	443	437	443

(أ) أختبر معنوية الفروق بين البكرات ؟

(ب) اختبر معنوية الفروق بين الدهانات ؟

٤٥- في دراسة لمقارنة ثلاثة مستويات من digitalis على مستوى الكالسيوم في عضلة قلب الكلب (الوصف الحقيقي للتجربة تم حذفه) حيث أخذت أنسجة القلب لأربعة حيوانات كل نسيج وزع عشوائيا على المعالجات الثلاثة وقد تم تقدير مستوى الكالسيوم في الأنسجة للمعالجات الثلاثة والبيانات في الجدول التالي :

المعالجة	مستويات الأنسجة			
	1	2	3	4
A	1342	1387	1549	1150
B	1881	1140	1296	1579
C	1608	1698	1029	1319

(أ) أجرِ تحليل التباين . هل هناك تفاوت في متوسطات المعالجات المختلفة ؟ استعمل مستوى معنوية $\alpha=0.05$.

(ب) هل هناك تفاوت في متوسطات العينات المختلفة (استعمل مستوى معنوية $\alpha=0.05$.

٤٦- فيما يلي درجات سمة الانبساطية لدي أربعة مجموعات وفي داخل كل مجموعة ذكور وإناث .

المجموعة	الجنس	
	ذكـور	إنـاث
المجموعة الأولى	5, 4, 3, 2, 6	3, 2, 4, 5, 6

←

المجموعة الثانية	7, 5, 2, 4, 3	7, 4, 5, 3, 2
المجموعة الثالثة	6, 7, 8, 9, 10	8, 8, 9, 77
المجموعة الرابعة	9, 18, 8, 7, 16	15, 8, 9, 8 ,8

(أ) المطلوب التحقق من صحة الفرض القائل .. لا توجد فروق معنوية في درجة السـمة الانبساطية بين المجموعات المختلفة .

(ب) لا توجد فروق معنوية في سمة الانبساطية بين الذكور والإناث .

(ج) لا يوجد تفاعل بين الجنس و المجاميع المختلفة .

٤٧- طبق اختبار للقلق على مجموعتين من الذكور والإناث في مجموعات عمريه مختلفـة وجاءت درجاتهم كما يلي :

المراحل العمرية	الجنس	
	ذكـور	إنـاث
الطفولـة	2,3 2, 4, 5	3, 4, 4, 2, 7
المراهقة	13, 15, 12, 8, 11	12, 6, 17, 7, 12
الشباب	5, 6, 6, 7, 4	5, 4, 5, 6, 7, 7

المطلوب :

(أ) هل توجد فروق بين الجنسين في القلق ؟

(ب) هل توجد فروق بين فئات العمر في القلق ؟

(ج) هل يوجد تداخل بين عامل الجنس والعمر على القلق ؟

٤٨- أجريت دراسة طبية على تأثير ثلاثة أنواع من الأدوية على سلوك مجموعتين مـن المرضى النفسيين (إكتئابيين – إنفصاميين) الدرجات التي حصلوا عليها في الجدول التالي :

نوع المرضى	نـوع الـدواء		
	1	2	3
أكتئابيين	4, 8, 0	10, 8, 6	8, 6, 4
أنفصاميين	4, 10, 6	4, 2, 0	15, 12, 9

(ب) أختبر معنوية الفروق بين الأدوية ؟

(ت) اختبر معنوية الفروق بين النوعين من المرضى ؟

(ث) هل هناك تفاعل بين نوع الدواء ونوع المرضى ؟

-٤٩- تم إجراء تجربة زراعية لدراسة تأثير الأنواع المختلفة وكذلك طرق الزراعة المختلفة (من حيث كثافة النباتات في مساحة معينة وهي كالتالي 10,20, 30, 40 (الف نبات لكل هكتار) على إنتاجية محصول الطماطم. البيانات في الجدول التالي :

كثافة النباتات				
النوع	10,000	20,000	30,000	40,000
H	10.5, 9.2, 7.9	12.8, 11.2, 13.3	12.1, 12.6, 14	10.8, 9.1, 12.5
Ife	8.1, 8.6, 10.1	12.7, 13.7, 11.5	14.4, 15.4, 13.7	11.3, 12.5, 14.5
P	16.1, 15.3, 17.5	16.6, 19.9, 18.5	20.8, 18, 21	18.4, 18.9, 17

المطلوب تحليل هذه البيانات للفروق المعنوية في المتوسطات بين الأنواع المختلفة من الطماطم وبين مستويات الكثافات المختلفة والتفاعل بين النوع والكثافة للنباتات عند مستوى معنوية $\alpha=0.05$.

-٥٠- في تجربة لدراسة تأثير الكميات المختلفة من Carbon fiber والكميات المختلفة من الرمل المضافة على عملية molding وذلك خلال صناعة الورقة ، تم الحصول على البيانات التالية :

	Carbon fiber		
الرمل المضاف	0.0	0.25	0.5
0.0	61.0 63.0	69 69	76 69
15	69 67	69 74	69 74
30	56 74	74 72	74 74

المطلوب تحليل هذه البيانات بأسلوب تحليل التباين عند مستوى معنوية $\alpha=0.01$.

-٥١- في تجربة زراعة لدراسة تأثير الأنواع المختلفة من البطاطا A, B, C وكذلك المناطق الجغرافية المختلفة على إنتاجية محصول البطاطا ، تم اختيار 9 قطع متساوية في المساحة في كل منطقة جغرافية ثم زراعة كل صنف من الأصناف الثلاثة في 3 قطع اختيرت عشوائياً والبيانات في الجدول التالي :

المنطقة الجغرافية	أنواع البطاطا		
	A	B	C
1	14 18 11	19 23 17	21 16 13
2	16 9 12	23 17 21	18 20 21
3	8 5 12	11 16 9	9 6 7
4	4 7 10	20 15 13	18 14 11

أوجد جدول تحليل التباين ثم أختبر التأثيرات والتفاعلات عند مستوى $\alpha=0.05$.

-٥٢- قام مهندس بقياس التيار (μ A) والضروري لإنتاج مستوى معين من الوضـــوح (brightnes) بصمام التليفزيون وذلك لأنواع مختلفة من الزجاج وأنواع مختلفـة مـن الفوسفور والبيانات في الجدول التالي :

		نوع الفسفور		
نوع الزجاج	1	1	2	3
	2	280, 290, 285	300, 310, 295	270, 285, 290
	3	230, 235, 240	260, 235, 240	220, 225, 230

والمطلوب تحليل هذه البيانات بأسلوب تحليل التباين .

-٥٣- اختيرت خمس عينات عشوائية من الطحالب plankton من المناطق A, B, C على بحيرة خلال شهر مايو . وقد كررت التجربة مرة أخـري في شـهر أغسـطس. البيانات (مقاسه thousands of plankton) معطاة في الجدول التالي :

(أ) أختبر معنوية الفروق بين المواقع . ، (ب) أختبر معنوية الفروق بين أوقات الجمع .

(ج) أختبر التفاعل بين الموقع وميعاد الجمع .

الموقع	ميعاد الجمع	
	أغسطس	مايو
A	97, 102, 109, 99, 101	107, 112, 118, 108, 111
B	105, 110, 115, 110	110, 115, 119, 110, 112

-٥٤- تعطى البيانات التالية الحموضة الكلية لعينات من ثلاثة أنواع من الفحم تم تحليلـها باستخدام تركيزات مختلفة من مركب ethanoloic Na OH .

المركب	نوع الفحم		
	Morwell	Yallourn	Maddingley
0.404 N	8.27, 8.17	8.66, 8.61	8.14, 7.96
0.626 N	8.03, 8.21	8.42, 8.28	8.02, 7.89
0.786 N	8.60, 8.20	8.61, 8.76	8.13, 8.07

أوجد جدول تحليل التباين وأختبر التفاعل ثم أختبر التأثيرات الأساسية لكل عامل (عنـد مستوي معنوية (0.0) .

-٥٥- تعطى البيانات التالية نتائج أربعة إختبارات في أربعة مقررات في جامعة ما لخمسـة من الطلبة.

الطالب	اللغة الإنجليزية		اللغة الفرنسية		الإحياء		الرياضيات	
1	50	57	72	80	86	80	89	62
	71	64	76	77	90	77	78	79
2	89	94	81	35	80	92	78	95
	66	87	79	67	61	92	55	67
3	73	46	90	95	77	70	66	65
	58	81	59	91	83	72	50	88
4	75	48	42	51	55	48	85	59
	25	75	41	31	12	55	63	70
5	93	93	95	80	80	76	89	76
	62	75	97	95	86	80	86	94

أستخدم مستوى معنوية α=0.05 لإختبار الفروض التالية :

(أ) هل هناك تفاوت في صعوبة المقررات الأربعة ؟

(ب) هل هناك تفاوت في مقدرات الطلبة . استعمل مستوى دلالة 0.05 .

(ت) هل يوجد تداخل بين الطلبة والمقررات ؟

-٥٦- في تجربة لمقارنة أربعة أنواع من الإطارات وثلاثة أنواع من الطرق علـى عـدد آلاف الكيلو مترات التي قطعت قبل تلف الإطار تم الحصول على البيانات التالية :

الطرق	نوع الإطار			
	A	B	C	D
1	9.9	8.9	7.5	9.4
	10.5	7.9	7.3	9.7
	11.3	7.6	7.2	10.6
	10.6	7.4	8.3	9.5

2	8.3 7.2 7.8 7.1	01.0 9.9 9.6 10.5	8.6 8.4 8.5 9.2	10.2 10.3 9.3 9.2
3	9.6 8.2 8.1 9.2	8.6 8.3 8.5 8.1	6.8 6.8 7.3 7.2	9.2 9.5 8.6 8.2

أستخدم مستوى معنوية $\alpha=0.01$ لاختبار الفروض التالية :

(أ) هل هناك فروق معنوية بين الأنواع ؟

(ب) هل هناك فروق معنوية بين الطرق ؟

(ث) هل يوجد تفاعل بين الأنواع والطرق ؟

-٥٧- في دراسة لتقدير تأثير نوعين من الإعلانات على كمية المبيعات لثلاثة أنواع من الكيك تم تسجيل المبيعات لكل نوع بعد الإعلان أ ثم بعد الإعلانات ب . ثم كررت التجربة ثلاثة مرات لكل إعلان . نتائج التجربة في الجدول التالي

الإعـــلان			
		أ	ب
الكيك	A	571, 563, 559	1091, 1076, 1065
	B	553, 570, 550	1027, 1072, 999
	C	576, 453, 591	1065, 1077, 1051

(أ) أوجد جدول تحليل التباين ؟

(ب) أختبر معنويــة الفــروق بــين الأنــواع المختلفــة عنــد مســتوى معنويــة $\alpha=0.05$.

(ج) أختبر معنوية الفروق بين الإعلانات عند مستوى معنوية $\alpha=0.05$.

(د) هل هناك تفاعل بين أنواع الكيك والإعلانات عند مستوى معنوية $\alpha=0.05$.

-٥٨- أجرى اختبار على مجموعة من الأطفال وذلك لبيان تأثير الدافـــع الشـــخصي أو تشجيع الوالدين على ذكاء الأطفال وكانت درجات الذكاء كالآتي :

تشجيع الوالدين	دافــع شخصــي	
	عالــي	منخفض
عالــي	121, 116, 110, 103, 104, 115, 117, 116, 118, 119	112, 101, 111, 102, 79, 74, 91, 91, 80, 89

←

منخفض	116, 93, 94, 93, 98, 95, 98, 97, 95, 94	69, 65, 91, 91, 65, 70, 71, 72, 72, 73

المطلوب تحليل هذه البيانات بأسلوب تحليل التباين عند مستوي معنوية α=0.05 .

الفصل الثانـي عشر

الاختبارات اللامعلميه

Nonparametric Tests

(١٢-١) مقدمـــة **Introduction**

تعتمد الطرق المستخدمة في اختبارات الفروض وتحليل التباين وتحليـــل الانحـــدار وتحليـــل الارتباط (الطرق المعلمية parametric methods)، والتي سبق مناقشـــتها في الفصـــول السابقة، على عدد من الفروض . فعلي سبيل المثال يشترط في الاختبار الذي يخـــص متوســـط مجتمع (اختبار t) أن المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً وذلك عندمـــا يكـــون تباين المجتمع غير معروف وحجم العينة صغير . أيضا في تحليل التباين يشترط أن المجتمعات الـــتي اختيرت منها العينات تتبع توزيعات طبيعية وتبايناتها متساوية . الاختبارات السابقة سوف تكـــون غير مجدية إذا لم تتحقق الشروط الخاصة بها ، وفي هذه الحالة تكون الاختبارات اللامعلمية هـــي البديل . عموماً تستخدم الاختبارات اللامعلمية في الأحوال الآتية :

(أ) عندما تكون الشروط اللازمة للاختبار المعلمي غير مستوفاة .

(ب) عندما يدور فرض العدم والفرض البديل علي أشياء وصفية وليست علي معلمة مجتمـــع كما في اختبار جودة التوفيق والذي سوف تتناوله في البند التالي . أيضا عند الرغبـــة في عمل مقارنة بين مجتمعين أو اكثر وذلك بالاعتماد على عينات عشوائية مختارة من هـــذه المجتمعات دون التعرف على التوزيعات الاحتمالية أو التعرض لها .

(ج) عندما يكون مقياس البيانات وصفي أو ترتيبي .

(د) عند الحاجة للوصول إلى قرار سريع بدون اســـتخدام الآلات الحاســـبة أو الحاســـبات الإلكترونية .

بعض مميزات الطرق اللامعلمية

(أ) فرص عدم الدقة قليلة عند تطبيقها وذلك لاعتمادها على اقل قدر من الفروض .

(ب) لبعض الطرق اللامعلمية ، العمليات تتم بسرعة لذلك فإنها توفر الوقت وخصوصاً عنـــد عدم توفر الآلات الحاسبة .

(ج) تلائم كثير من الباحثين في مجال علم النفس والاجتماع لأن معظـــم البيانـــات الذيـــن يتعاملون معها تكون من النوع الوصفي أو الترتيبي .

(د) تلائم الباحثين الذين لديهم أدنى معلومات في مجال الرياضيات والإحصاء وذلك لســـهولة المفاهيم والطرق اللامعلمية .

بعض عيوب الطرق اللامعلميه

(أ) ولان العمليات المستخدمة في معظم الاختبارات اللامعلمية بسيطة وسريعة فإنها تؤدى إلى فقد في المعلومات الموجودة في البيانات كما هو الحال عند تحويل البيانات إلى رتب ، وهذا يؤدى إلى فقد كبير في الدقة .

(ب) بعض الطرق اللامعلمية تكون معقدة .

(١٢-٢) اختبار مربع كاي لجودة التوفيق

Chi-square Goodness-of-fit Test

أن عملية التعرف على التوزيع الاحتمالي للمجتمع الذي اختيرت منه عينة عشوائية من الشروط اللازمة لتطبيق بعض الاختبارات كما أوضحنا في البند السابق . يستخدم اختبار مربع كاي لجودة التوفيق لاختبار ما إذا كانت مشاهدات عينة عشوائية تم اختيارها من مجتمع له توزيع احتمالي معين .

تتكون البيانات اللازمة للاختبار من عينه عشوائية من الحجم **n** من المشاهدات المستقلة . تصنف المشاهدات إلى **k** من الفئات الشاملة المانعة (الخلايا **cell**) كما هو موضح في جدول (١٢-١). عدد المشاهدات التي تقع في فئة معطاة تسمى التكرار المشاهد **observed frequency** لهذه الفئة حيث O_i تمثل عدد المشاهدات في الفئة رقم **i** ، $i = 1,2,\ldots, k$.

جدول (١٢-١)

الفئة	1	2	...	i	...	k
التكرار المشاهد	O_1	O_2	...	O_i	...	O_k

الفئات قد تكون أسمية (وصفية) أو عددية . على سبيل المثال قد تنتمي مشاهدات العينة إلى واحدة من الفئتين الاسميتين ذكر وأنثى. أيضا إذا كان المتغير موضع الدراسة هو العمر فإن مشاهدات العينة قد تتبع واحدة من الفئات العمريه التالية :

$$<15, 15\text{-}24, 25\text{-}34, 35\text{-}44, 45\text{-}54, \geq 55$$

لإجراء الاختبار نعرف الاحتمال ، سوف يرمز له بالرمز P_i ، بأن مشاهدة اختيرت عشوائياً من المجتمع النظري (المفترض) سوف تقع في الفئة رقم **i** وعلى ذلك نعين الاحتمالات $P_1, P_2, \ldots, P_k$ للفئات $i = 1,2,\ldots,k$ على التوالي . عندما يكون فرض العدم صحيح ، فإنه يمكن حساب التكرارات المتوقعة لكل فئة أي $n_1P_1, n_2P_2, \ldots, n_kP_k$ للفئات $i = 1,2,\ldots,k$ على التوالي والتي سوف يرمز لها بالرمز $E_1, E_2, \ldots, E_k$. فرض العدم والفرض البديل سوف يكونان على الشكل :

H_0 : العينة اختيرت من مجتمع يتبع توزيع احتمالي معين .

H_1 : العينة اختيرت من مجتمع لا يتبع هذا التوزيع الاحتمالي المعين .

للعينات الكبيرة وبفرض أن H_0 صحيح فإن :

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

قيمة لمتغير عشوائي X^2 تقريباً يتبع توزيع χ^2 بدرجات حرية **k-1** . لمستوى معنوية α فإن منطقة الرفض $X^2 > \chi^2_\alpha$ حيث أن χ^2_α تستخرج من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية **k-1** . إذا وقعت χ^2 في منطقة الرفض نرفض H_0 . يكون التقريب مقبولاً إذا كان عدد مشاهدات العينة أكبر من **50** والتكرار المتوقع المناظر لكل فئة لا يقل عن **5** . في بعض الأحيان تستخدم مشاهدات العينة في تقدير معلمة أو اكثر من معالم المجتمع ثم يستخدم هذا التقدير في حساب التكرارات المتوقعة . فإذا كان عدد المعالم المقدرة هو **m** فإن درجات الحرية في هذه الحالة تصبح **(k-m-1)** .

مثال (١٢-١) يقوم المسئول في مصنع لإنتاج معاطف السيدات العالية الجودة بإرسال كل معطف منتج إلى واحد من مراكز مراقبة الجودة. يعطى جدول (١٢-٢) التوزيع التكراري لعدد المعاطف المرفوضة لعينة عشوائية من الحجم **n=100** . هل يمكن القول أن عدد المعاطف المرفوضة واحدة لكل مركز ؟

جدول (١٢-٢)

مركز الفحص	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2 / E_i$
1	15	10	5	25	2.5
2	3	10	-7	49	4.9
3	7	10	-3	9	0.9
4	8	10	-2	4	0.4
5	5	10	-5	25	2.5
6	12	10	+2	4	0.4
7	11	10	+1	1	0.1
8	13	10	+3	9	0.9
9	9	10	-1	1	0.1
10	17	10	+7	49	4.9
المجموع	100	100			17.6

الحل.

H_0 : المراقبون يرفضون نفس العدد من المعاطف (العينة تم اختيارها من مجتمع يتبـــع التوزيـــع المنتظم).

H_1 : المراقبون لا يرفضون نفس العدد من المعاطف (العينة تم اختيارها مـــن مجتمـــع لا يتبـــع التوزيع المنتظم).

تحت فرض العدم فإن $P_i = \frac{1}{10}, i = 1,2,...,10$. التكرارات المتوقعة في جـدول (١٢-٢) تم حسابها من الصيغة $E_i = nP_i = (100)(\frac{1}{10}) = 10$ حيث $i = 1,2,...,10$. لمســتوى معنوية $\alpha=0.05$ فإن $\chi^2_{.05} = 16.919$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) عنـد درجات حرية $k-1 = 10 - 1 = 9$. منطقة الرفض $X^2 > 16.919$. وبما أن χ^2 تقـــع في منطقة الرفض نرفض H_0 .

مثال (١٢-٢) يعطي الجدول (١٢-٣) عدد الوفيات في الأسبوع x والتي وقعت في 10 مدن خلال 200 أسبوع والناتجة من حوادث السيارات .

جدول (١٢-٣)

x	0	1	2	3	4
التكرارات المشاهدة	109	65	22	3	1

هل النتائج المعطاة في جدول (١٢-٣) تتفق مع الفرض القائل أن عدد الوفيات في الأســـبوع والناتجة من حوادث السيارات تتبع توزيع بواسون ؟ استخدم مستوى معنوية $\alpha=0.05$.

الحل .

H_0: عدد الوفيات في الأسبوع والناتجة من حوادث السيارات تتبع توزيع بواسون .

H_1 : عدد الوفيات في الأسبوع والناتجة من حوادث السيارات لا تتبع توزيع بواسون .

بفرض أن H_0 صحيح فإن الدالة الاحتمالية لتوزيع بواسون هي :

$$P(X = i) = \frac{\bar{e}^{\mu}\mu^{i}}{i!}, i = 1,2,...$$

وحيث أن μ غير محددة من فرض العدم فإنه يمكن تقديرها من مشاهدات العينة وأفضل تقديـــر للمعلمة μ هو الوسيط الحسابي للمشاهدات والذي يتم حسابه من جدول (١٢-٣) حيــث أن :

$$\hat{\mu} = \overline{x} = \frac{\Sigma x_i O_i}{\Sigma O_i} = \frac{122}{200} = 0.61.$$

باستبدال μ في توزيع بواسون بالقيمة $\hat{\mu}$ نحصل على :

$$P_i = P(X = i) = \frac{(0.61)^i \ \bar{e}^{0.61}}{i!}, i = 0,1,2,3,4 .$$

التكرارات المتوقعـــة لعـــدد الوفيـــات في الأســـبوع والناتجـــة مـــن حـــوادث الســـيارات ، $E_i = nP_i, i = 0,1,2,3,4$ ، معطاة في جدول (١٢-٤).

جدول (١٢-٤)

x	التكرارات المتوقعـــة
0	$200.P(X=0) = 200\frac{(0.61)^0 \ \bar{e}^{.61}}{0!} = 108.7$
1	$200.P(X=1) = 200\frac{(0.61)^1 \ \bar{e}^{.61}}{1!} = 66.3$
2	$200.P(X=2) = 200\frac{(0.61)^2 \ \bar{e}^{.61}}{2!} = 20.2$
3	$200.P(X=3) = 200\frac{(0.61)^3 \ \bar{e}^{.61}}{3!} = 4.1$
4	$200.P(X=4) = 200\frac{(0.61)^4 \ \bar{e}^{.61}}{4!} = 0.7$

يعطي الجدول (١٢-٥) التكرارات المشاهدة والتكرارات المتوقعة بعد تقريبها .

جدول (١٢-٥)

x	0	1	2	3	4	5 أو أكثر	المجموع
التكرارات المتوقعة	109	66	20	4	1	0	200
التكرارات المشاهدة	109	65	22	3	1	0	200

بالنظر إلى التكرارات المتوقعة في جدول (١٢-٥) نجد أن هناك تكرارين كل منهما أقل من 5 وهما التكرار الرابع والتكرار الخامس وبذلك لا يصلح تطبيق اختبار مربع كاى إلا بعد التغلب على هذه المشكلة وذلك بدمج التكرارين كما هو معطى في جدول (١٢-٦) :

جدول (١٢-٦)

x	0	1	2	أكبر أو يساوى 3
التكرارات المتوقعة	109	66	20	5
التكرارات المشاهدة	109	65	22	4

من جدول (١٢-٦) فإن :

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 0.41515.$$

لمستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{.05} = 5.992$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية k-m-1=4-1-1=2 . منطقة الرفض $\chi^2 > 5.992$. وبما أن χ^2 تقع في منطقة القبول نقبل H_0 .

مثال (١٢-٣) يعطى جدول (١٢-٧) التوزيع التكراري لأطوال 40 بطارية. هل البيانات في جدول (١٢-٧) تتفق مع القول أن التوزيع الطبيعي يعطى توفيق جيد لتوزيع أعمار البطاريات وذلك عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢-٧)

الحدود الفعلية للفئة	التكرارات المشاهدة	التكرارات المتوقعة
1.45 - 1.95	2	0.6
1.95 – 2.45	1	2.6
2.45 – 2.95	4	6.8
2.95 – 3.45	15	10.7
3.45 – 3.95	10	10.3
3.95 – 4.45	5	6.1
4.45 – 4.95	3	2.2
المجموع	40	

الحل .

H_0 : العينة اختيرت من مجتمع يتبع التوزيع الطبيعي .

H_1 : العينة اختيرت من مجتمع لا يتبع التوزيع الطبيعي .

التكرارات المتوقعة في جدول (١٢-٧) تم حسابها من منحنى طبيعـــي لــه نفــس المتوســط والانحــراف المعيــاري لمشــاهدات العينــة المعطـــــاة في جـــدول (١٢-٧) حيـــث $\bar{x} = 3.4125$, $s = 0.697$ أي أننا سوف نســتخدم $\bar{x}$, s (بـــدلا مـــن μ , σ الخاصة بالمجتمع الذي اختيرت منه العينة) في حساب قيم z. على سبيل المثال للفئة الرابعة فإن :

$$z_1 = \frac{2.95 - 3.4125}{0.697} = -0.67,$$

$$z_2 = \frac{3.45 - 3.4125}{0.697} = 0.054.$$

ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن المساحة بين $z_1 = -0.67$, $z_2 = 0.054$ هي :

$$P(-0.67 < Z < 0.054)$$

$$= P(0 < Z < 0.67) + P(0 < Z < 0.054)$$

$$= 0.2486 + 0.0199 = 0.2685.$$

وعلى ذلك فإن التكرارات المتوقعة للفئة الرابعة هي :

$$E_4 = (0.2685)(40) = 10.7$$

حيث أن $\sum_{i=1}^{k} O_i = 40$.التكرار المتوقع للفئة الأولي تم الحصول عليه باستخدام المساحة الكليـة تحت المنحنى الطبيعي على يسار القيمة 1.95 (الحد الأعلى الفعلي للفئة الأولى). للفئة الأخيرة استخدمت المساحة الكلية تحت المنحنى الطبيعي على يمين القيمة 4.45 (الحد الأدنى الفعلي للفئة الأخيرة). التكرارات المتبقية تم حسابها بنفس الطريقة التي شرحناها للفئـــة الرابعــة . بدمـــج التكرارات المتوقعة للفئات التي تكراراتها المتوقعة أقل من 5 نحصل على جدول (١٢-٨) .

جدول (١٢-٨)

الحدود الفعلية	التكرارات المشاهدة	التكرارات المتوقعة
1.45 – 2.95	7	10
2.95 – 3.45	15	10.7
3.45 – 3.95	10	10.3
3.95 – 4.95	8	8.3

من جدول (١٢- ٨) فإن :

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 2.648.$$

لمستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{.05} = 3.843$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $k-m-1=4-2-1=1$. منطقة الرفض $\chi^2 > 3.843$. وبما أن χ^2 تقع في منطقة القبول نقبل H_0.

(١٢-٣) اختبار مربع كاى للاستقلال The Chi-square Test of Independent

في كثير من الأحيان يرغب الباحث في التعرف عما إذا كانت هناك علاقة بين صفتين من صفات مجتمع ما. فعلى سبيل المثال قد يرغب مسئول التغذية في مدرسة ما في التعرف عما إذا كانت الحالة الغذائية للطالب لها علاقة بكفاءته التعليمية . أيضا قد يرغب باحث في مجال الوراثة في التعرف عما إذا كانت هناك علاقة بين لون الشعر ولون العينين ... الخ .

لإجراء الاختبار نختار عينة عشوائية من الحجم n من المجتمع موضع الدراسة .تصنف مشاهدات هذه العينة حسب مستويات كل من الصفتين موضع الدراسة في جدول مزدوج يسمى جدول التوافق Contingency table . بفرض أن $A_1, A_2,..., A_k$ ترمز لمستويات الصفة A و $B_1, B_2,..., B_k$ ترمز لمستويات الصفة B فإن جدول التوافق يكون على الشكل الموضح في جدول (١٢-٩) ، حيث أن O_{ij} ترمز لعدد المشاهدات التي يتوفر فيها المستوى A_i من الصفة A و المستوى B_j من الصفة B حيث $i = 1,2,...,r$ و $j = 1,2,...,c$.
أيضا $n_{i.}$ ترمز لعدد المشاهدات التي يتوفر فيها المستوى A_i من الصفة A أي أن $n_{i.} = \sum_{j=1}^{c} O_{ij}$. أيضا $n_{.j}$ ترمز لعدد المشاهدات التي يتوفر فيها المستوى B_j من الصفة B أي أن $n_{j.} = \sum_{i=1}^{r} O_{ij}$. وعلى ذلك :

$$\sum_{j=1}^{c} n_{.j} = \sum_{i=1}^{r} n_{i.} = \sum_{j=1}^{c} \sum_{i=1}^{r} O_{ij}.$$

يحتوي جدول التوافق على خانات (خلايا) عددها (r x c) خلية.

جدول (١٢ -٩)

	B_1 B_2 .... B_c	المجموع
A_1	O_{11} O_{12} ... O_{1c}	$n_{1.}$
A_2	O_{21} O_{22} ... O_{2c}	$n_{2.}$
⋮	⋮ ⋮ ⋮ ⋮	⋮
A_r	O_{r1} O_{r2} ... O_{rc}	$n_{r.}$
	$n_{.1}$ $n_{.2}$ ... $n_{.c}$	n

فرض العدم والفرض البديل سوف يكونان علي الشكل :

H_0 : المتغيرين مستقلين .

H_1 : المتغيرين غير مستقلين .

يعتمد اختبار مربع كاي للاستقلال علي مقارنة التكرارات المشاهدة بالتكرارات المتوقعة في كل خلية عندما H_0 صحيح . إذا كان $P(A_i)$ يرمز لاحتمال أن يتوفر لمشاهدة ما المستوى A_i من الصفة A وإذا كان $P(B_j)$ يرمز لاحتمال أن يتوفر لمشاهدة ما المستوى B_j من الصفة B وإذا كان P_{ij} يرمز لاحتمال أن يتوفر لمشاهدة ما المستوى A_i من الصفة A و المستوى B_j من الصفة B فإن :

$$P_{ij} = P(A_i \cap B_j)$$

وفي حالة الاستقلال بين الصفتين A , B (تحت فرض العدم) فإن :

$$P_{ij} = P(A_i) \cdot P(B_j).$$

يمكن الحصول علي تقدير للاحتمال P_{ij} كالتالي :

$$P'_{ij} = \left(\frac{n_{i.}}{n}\right)\left(\frac{n_{.j}}{n}\right)$$

وعلي ذلك يمكن حساب التكرارات المتوقعة كالتالي :

$$E_{ij} = n\left(\frac{n_{i.}}{n}\right)\left(\frac{n_{.j}}{n}\right)$$

$$= \frac{n_{i.} \quad n_{.j}}{n} \quad ,i = 1,2,\ldots,r, \quad ; j = 1,2,\ldots,c.$$

بافتراض أن H_0 صحيح فإن :

$$\chi^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

قيمة لمتغير عشوائي X^2 تقريبا يتبع توزيع χ^2 بدرجات حرية $(r-1)(c-1)$ حيث r عدد الصفوف و c عدد الأعمدة في جدول التوافق . لمستوى معنوية α فإن منطقة الرفض $X^2 > \chi^2_\alpha$ حيث χ^2_α تستخرج من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $(r-1)(c-1)$. إذا وقعت χ^2 في منطقة الرفض نرفض H_0 .

مثال (١٢-٤) يعتقد الأطباء أن عدد ساعات النوم لسيدة لديها أطفال يختلف عن عدد ساعات النوم قبل إنجابها. بفرض إنه تم سؤال 60 سيدة لديها أطفال وسجلت البيانات في جدول

(١٢-١٠). ما هو الاستدلال الذي يمكن الحصول عليه من هذه البيانات ؟ عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢-١٠)

عدد الأطفال	النوم الحالي بالمقارنة قبل الإنجاب			المجموع
	أقل	نفسه	أحسن	
1	25	5	0	30
2	10	4	1	15
3 أو أكثر	5	7	3	15
المجموع	40	16	4	60

الحل .

H_0 : عدد ساعات النوم وعدد الأطفال مستقلين .

H_1 : عدد ساعات النوم وعدد الأطفال غير مستقلين .

التكرارات المتوقعة معطاة في جدول (١٢-١١) .

جدول (١٢-١١)

عدد الأطفال	النوم الحالي بالمقارنة قبل الإنجاب		
	أقل	نفسه	أحسن
1	20	8	2
2	10	4	1
3 أو أكثر	10	4	1

التكرارات المشاهدة والتكرارات المتوقعة بعد دمج بعض التكرارات معطاة في جدول (١٢-١٢) و جدول (١٢-١٣) على التوالي وذلك حتى يتحقق الشرط أن عدد التكرارات المتوقعة في كل خلية لا يقل عن 5 .

جدول (١٢-١٢)

عدد الأطفال	اقل	نفسه أو أحسن
1	25	5 + 0 = 5
2	10	4 + 1 = 5
3 أو أكثر	5	7 + 3 = 10

من جدول (١٢-١٢) وجدول (١٢-١٣) يمكن حساب :

$$\chi^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 11.25.$$

لمستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{.05} = 5.992$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $2x1 = 2$. منطقة الرفض $X^2 > 5.992$. وبما أن χ^2 تقع في منطقة الرفض فإننا نرفض H_0 .

جدول (١٢-١٣)

عدد الأطفال	اقل	نفسه أو أحسن
1	20	10
2	10	5
3 أو أكثر	10	5

مثال (١٢- ٥) لدراسة العلاقة بين لون شعر الزوج والزوجة قام باحث بإختيار عينة عشوائية من الحجم 500 (زوج وزوجة) وتم سؤالهم والبيانات في جدول (١٢-١٤).

جدول (١٢-١٤)

الزوجة	الزوج				المجموع
	أحمر	أصفر	أسود	بنى	
أحمر	10	10	10	20	50
اصفر	10	40	50	50	150
اسود	13	25	60	52	150
بنى	17	25	30	78	150
المجموع	50	100	150	200	500

المطلوب اختبار ما إذا كانت هناك علاقة بين لون شعر الزوج والزوجة أم لا ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل .

H_0 : لون شعر الزوج ولون شعر الزوجة مستقلين.

H_1 : لون شعر الزوج ولون شعر الزوجة غير مستقلين.

التكرارات المتوقعة معطاة في جدول (١٢-١٥) . من جدول (١٢-١٤) و جدول (١٢-١٥) فإن :

$$\chi^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$= 32.56.$$

جدول (١٢- ١٥)

الزوجة	الزوج				المجموع
	أحمر	أصفر	أسود	بنى	
أحمر	5	10	15	20	50
اصفر	15	30	45	60	150
اسود	15	30	45	60	150
بنى	15	30	45	60	150
المجموع	50	100	150	200	500

لمستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{.05} = 16.919$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $3 \times 3=9$. منطقة الرفض $X^2 > 16.919$ وبما أن χ^2 المحسوبة تقع في منطقة الرفض فإننا نرفض H_0 .

إذا كان لكل من الصفتين A,B مستويان فقط فإن الجدول الناتج يتكون من صفين وعمودين (أى أربع خلايا). يسمى الجدول الناتج جدول الاقتران (2x 2) . جدول (١٢- ١٦) يمثل جدول اقتران. عدد درجات الحرية التي ترتبط بجدول الاقتران سوف تساوى الواحد الصحيح.

جدول (١٢-١٦)

الصفة الأولى	الصفة الثانية		
	B_1	B_2	
A_1	a	b	a + b
A_2	c	d	c + d
	a+c	b+d	n

يمكن استخدام صيغة بسيطة لحساب قيمة χ^2 كالتالي :

$$\chi^2 = \frac{n(ad-bc)^2}{(a+c)(b+d)(c+d)(a+b)}.$$

مثال (١٢-٦) لدراسة العلاقة بين النوم ليلا والتدخين اختيرت عينة عشوائية من 56 شخصا والبيانات معطاة في جدول (١٢-١٧) .

جدول (١٢-١٧)

التدخين	النوم		المجموع
	نعم	لا	
نعم	20	16	36
لا	6	14	20
المجموع	26	30	56

المطلوب اختبار ما إذا كانت هناك علاقة بين النوم ليلا والتدخين وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل .

H_0 : المتغيرين مستقلين .

H_1 : المتغيرين غير مستقلين .

من جدول (١٢-١٧) فإن :

$$\chi^2 = \frac{n(ad-bc)^2}{(a+c)(b+d)(c+d)(a+b)}$$

$$= \frac{56[(20)(14)-(16)(6)]^2}{(26)(30)(20)(36)} = 3.376.$$

لمستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{.05} = 3.843$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية واحدة . منطقة الرفض $X^2 > 3.843$. وبما أن χ^2 تقع في منطقة القبول نقبل H_0 .

سبق أن ذكرنا أن التكرارات المتوقعة في كل خلية يجب أن لا يقل عن 5 وإذا حدث وكان أحد التكرارات المتوقعة أقل من 5 فإننا نقوم بدمج التكرارات . وعلي أى حال فإن هذه الطريقة لا تستخدم في حالة جدول الاقتران . وقد أقترح Yates (1934) تصحيحا يستخدم في

حالة ما إذا كان أحد التكرارات المتوقعة أقل من 5 . وباستخدام التصحيح يصبح قيمة الإحصاء الذي يعتمد عليه قرارانا هو :

$$\chi^2 = \frac{n(|ad - bc| - \frac{n}{2})^2}{(a+c)(b+d)(c+d)(a+b)}$$

بتطبيق تصحيح Yates علي البيانات في جدول (١٢-١٧) فإن قيمة الإحصاء تصبح :

$$\chi^2 = \frac{56[|(20)(14) - (16)(6)| - \frac{56}{2}]^2}{(26)(30)(20)(36)}$$

$$= 2.427.$$

لمستوى معنوية $\alpha = 0.05$ فإننا نحصل إلى نفس الاستنتاج الذي حصلنا عليه بدون تصحيح ، أى إننا نقبل H_0 .

(١٢ -٤) اختبار مربع كاي للتجانس

The Chi-square Test of Homogeneity

بفرض أن لدينا مجتمعات عددها r وجميعها متماثلة من حيث التصنيف وبفرض أن c هـي عدد فئات التصنيف في كل مجتمع . بفرض أن $P_{j|i}$ يرمز لنسبة مشاهدات المجتمع رقم i الـتي تقع في الفئة رقم j . يمكن تمثيل هذه المجتمعات بالجدول (١٢-١٨) .

جدول (١٢- ١٨)

المجتمع	فئات التصنيف 1 2 ... j ... c	
1	$P_{1\|1}$ $P_{2\|1}$ ... $P_{j\|1}$ ... $P_{c\|1}$	1
2	$P_{1\|2}$ $P_{2\|2}$ ... $P_{j\|2}$ ... $P_{c\|2}$	1
⋮		⋮
i	$P_{1\|i}$ $P_{2\|i}$ ... $P_{j\|i}$ ... $P_{c\|i}$	1
⋮		⋮
r	$P_{1\|r}$ $P_{2\|r}$ ... $P_{j\|r}$ ... $P_{c\|r}$	1
	P_1 P_2 .... P_j ... P_c	1

عندما تكون النسب $P_{j|i}$ مجهولة فإننا نرغب في معرفة ما إذا كانت المجتمعات التي عددها r متجانسة أي إننا نرغب في اختبار فرض العدم :

$$H_0 : P_{j|1} = P_{j|2} = ... = P_{j|r} = P_j$$

$$; j = 1,2,....c.$$

لإجراء الاختبار فإننا نختار عينات عشوائية عددها r واحدة من كل مجتمع وأحجامها هي $n_1, n_2, ..., n_r$ على أن تكون العينات العشوائية مستقلة عن بعضها البعض. بفحص مشاهدات هذه العينات ووضع كل مشاهدة حسب تصنيفها نحصل على الجدول (١٢-١٩).

جدول (١٢- ١٩)

المجتمع	1 2 ... j ... c	
1	O_{11} O_{12} ... O_{1j} ... O_{1c}	n_1
2	O_{21} O_{22} ... O_{2j} ... O_{2c}	n_2
⋮		⋮
i	O_{i1} O_{i2} ... O_{ij} O_{ic}	n_i
⋮		⋮
r	O_{r1} O_{r2} ... O_{rj} ... O_{rc}	n_r
المجموع	$n_{.1}$ $n_{.2}$ ... $n_{.j}$ ... $n_{.c}$	n

حيث $n_i = \sum_{j=1}^{c} O_{ij}$ و $n_{.j} = \sum_{i=1}^{r} O_{ij}$ و $n = \sum_{i=1}^{r} n_i = \sum_{j=1}^{c} n_{.j}$.

إذا كان فرض العدم صحيح وحيث أن النسب مجهولة فإننا نقوم بتقدير P_j حيث أن :

$$P'_j = \frac{n_{.j}}{n},$$

$$E_{ij} = n_{i.} \; P'_j = n_{i.}\left(\frac{n_{.j}}{n}\right),$$

وعلى ذلك فإن :

$$\chi^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

هي قيمة للمتغير العشوائي X^2 الذي تقريبا يتبع توزيع χ^2 بدرجات حرية $(r-1)(c-1)$. لاختبار فرض العدم عند مستوى معنوية α نتبع الخطوات التي استخدمناها في اختبار مربع كاي للاستقلال .

مثال (١٢-٧) قامت شركة للمياه الغازية بدراسة لمعرفة ما إذا كان هناك اختلاف بين شرائح مختلفة من المجتمع من ناحية التفضيل لثلاثة أنواع من المشروبات . استخدمت لهذه الدراسة أربع عينات مستقلة والنتائج معطاة في جدول (١٢-٢٠) . استخدم اختبار مربع كاى للتجانس لاختبار فرض العدم :

H_0 : المجتمعات الأربعة متساويين في تفضيل المشروب.

ضد الفرض البديل :

H_1 : المجتمعات الأربعة غير متساويين في تفضيل المشروب.

وذلك عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢-٢٠)

العينات الأربعة	نوع المشروب			المجموع
	A	B	C	
ربات البيوت	75	20	5	100
رجال الأعمال	50	130	20	200
عمال	5	25	17	47
طلبة	100	100	100	300
المجموع	230	275	142	647

الحل . التكرارات المتوقعة تم حسابها في جدول (١٢-٢١).

جدول (١٢-٢١)

العينات الأربعة	نوع المشروب			المجموع
	A	B	C	
ربات البيوت	35.55	42.50	21.95	100
رجال الأعمال	71.10	85.01	43.89	200
عمال	16.71	19.98	10.32	47.01
طلبة	106.65	127.51	65.84	300
المجموع	230.01	275	142	647.01

نحسب قيمة الإحصاء من الصيغة التالية :

$$\chi^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$
$$= 149.72.$$

لمستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{.05} = 12.592$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) عند درجات حرية $r \times c = 3 \times 2 = 6$. منطقة الرفض $X^2 > 12.592$. وبما أن χ^2 تقع في منطقة الرفض نرفض H_0 .

(١٢-٥) اختبار خاص بالاعتدال A Special Test for Normality

كثير من اختبارات الفروض التي تناولناها في الفصول السابقة تشترط أن مشاهدات العينة يجب أن تكون مختارة من مجتمع يتبع توزيعاً طبيعياً . فرض العدم والفرض البديل لاختبار الاعتدال يكونان على الشكل :

H_0 : توزيع المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً .

H_1 : توزيع المجتمع الذي اختيرت من العينة لا يتبع توزيعاً طبيعياً .

لإجراء الاختبار نختار عينة عشوائية من الحجم n من المشاهدات $x_1, x_2, \ldots, x_n$ وذلك من المجتمع موضع الدراسة. بفرض أن مشاهدات العينة تم ترتيبها من الأصغر إلى الأكبر (ترتيب تصاعديا) حيث $x_{(1)}, x_{(1)}, \ldots, x_{(n)}$ ترمز إلى المشاهدات بعد ترتيبها . يتم حساب القيم $z_{(1)}, z_{(1)}, \ldots, z_{(n)}$ من جدول التوزيع الطبيعي القياسي في ملحق (٣) والتي المساحة قبلها تساوى $[(i - 0.375)/(n + 0.25)]$. على سبيل المثال إذا كان حجم العينة $n = 20$ فإن $z_{(2)}$ تستخرج من جدول التوزيع الطبيعي القياسي في ملحق (٣) وهي القيمة التي المساحة قبلها $[(2-0.375)/20.25]=0.0802$. أي أن $z_{(2)} \simeq -1.4$. لأزواج القيم $(x_{(1)}, z_{(1)}), (x_{(2)}, z_{(2)}), \ldots, (x_{(n)}, z_{(n)})$ يتم حساب معامل الارتباط البسيط r. بفرض أن H_0 صحيح فإن r قيمة لمتغير عشوائي R_r له توزيع احتمالي . القيم الحرجة c_α من R_r معطاة في الجدول في ملحق (١٠) عند مستويات معنوية مختلفة . منطقة الرفض $R_r \leq c_\alpha$. إذا وقعت r في منطقة الرفض نرفض H_0 .

مثال (١٢-٨) يعطى الجدول (١٢-٢٢) مشاهدات لعينة عشوائية من الحجم n=20 . كل قيمة تمثل نسبة العرض إلى الطول لقطعة مربعة من جلد الثعبان مستخدمة في صناعة حقائب للسيدات .

جدول (١٢- ٢٢)

i	1	2	3	4	5	6	7	8	9	10
$x_{(i)}$	0.553	0.570	0.576	0.601	0.606	0.606	0.609	0.611	0.615	0.628
$z_{(i)}$	-1.87	-1.40	-1.13	-0.92	-0.66	-0.66	-0.45	-0.31	-0.19	-0.06
i	11	12	13	14	15	16	17	18	19	20
$x_{(i)}$	0.654	0.662	0.668	0.670	0.672	0.690	0.693	0.749	0.844	0.933
$z_{(i)}$	0.06	0.19	0.31	0.45	0.59	0.74	0.92	1.13	1.40	1.87

المطلوب اختبار فرض العدم :

H_0 : توزيع المجتمع الذي اختيرت منه العينة طبيعي .

ضد الفرض البديل :

H_1 : توزيع المجتمع الذي اختيرت منه العينة غير طبيعي .

وذلك عند مستوى معنوية $\alpha = 0.01$.

الحل . من البيانات في جدول (١٢-٢٢) فإن :

$$n = 20 \quad , \quad \Sigma\, x_{(i)} = 13.21 \quad , \quad \Sigma\, x_{(i)}^2 = 8.887812,$$

$$\Sigma\, z_{(i)} = 0.01 \quad , \quad \Sigma\, z_{(i)}^2 = 17.6039 \quad , \quad \Sigma\, x_{(i)} z_{(i)} = 1.5371$$

وعلى ذلك :

$$r = \frac{\Sigma\, x_{(i)} z_{(i)} - \dfrac{\Sigma\, x_{(i)} \Sigma\, z_{(i)}}{n}}{\sqrt{\left[\Sigma\, x_{(i)}^2 - \dfrac{(\Sigma\, x_{(i)})^2}{n}\right]\left[\Sigma\, z_{(i)}^2 - \dfrac{(\Sigma\, z_{(i)})^2}{n}\right]}}$$

$$= \frac{1.5371 - \dfrac{(13.21)(0.01)}{20}}{\sqrt{\left[8.887812 - \dfrac{(13.21)^2}{20}\right]\left[17.6039 - \dfrac{(0.01)^2}{20}\right]}}$$

$$= 0.905.$$

لمستوى معنوية $\alpha = 0.01$ فإن $c_{.as} = 0.929$ والمستخرجة من الجدول في ملحق (١٠) عند n = 20 . منطقة الرفض $R_r < 0.929$. بما أن r تقع في منطقة الرفض نرفض H_0 .

(١٢- ٦) اختبار الإشارة لعينة واحدة

The One – sample Sign Test

يستخدم اختبار الإشارة كبديل لاختبار t الخاص بمتوسط مجتمع وذلك عند عدم التحقق من أن المجتمع الذي اختيرت منة العينة يتبع توزيعاً طبيعياً وحجم العينة أقل من 30 .

تتكون البيانات اللازمة للتحليل من عينة عشوائية من الحجم n من المشـــاهدات $x_1,$ $x_2,\ldots,x_n$ والمختارة من مجتمع متصل وسيطه مجهول . في الحقيقة إذا كان التوزيع متماثل فــإن الوسيط يساوى الوسط الحسابي للمجتمع μ ويمكن استخدام اختبار الإشارة كاختبار للمتوسط . فرض العدم والفرض البديل سوف يكونان على الشكل :

$$H_0 : M = M_0,$$
$$H_0 : M > M_0.$$

لإجراء الاختبار نحسب القيمـــة k والــتي تمثـل عــدد الإشــارات الســالبة للفــروق $(x_i - M_0), i = 1,2,\ldots,n$ وإذا وجدت مشاهدة تساوى الوسيط نهملـــها ولا نأخذهـــا في الاعتبار . بفرض أن H_0 صحيح فإن k تمثل قيمة لمتغير عشوائي (إحصاء) K له توزيــع ذي الحدين بمعالم p=0.5 , n . لمستوى معنوية α نرفض H_0 إذا كان :

$$P(K \le k|\ n,0.5) \le \alpha .$$

للفرض البديل $H_0 : M < M_0$ نرفض H_0 إذا كان :

$$P(K \le k|\ n,0.5) \le \alpha .$$

حيث k تمثل عدد الإشارات الموجبة.

للفرض البديل $H_0 : M \ne M_0$ نرفض H_0 إذا كان :

$$P(K \le k|\ n,0.5) \le \frac{\alpha}{2}$$

حيــث k تمثــل عــدد الإشــارات الموجبــة أو الســالبة أيـــهما اقـــل للفـــروق $(x_i - M_0), i = 1,2,\ldots,n$.

مثال (١٢-٩) يعطى الجدول (١٢-٢٣) الدخول السنوية (بالآلاف الدولارات) لعينـــة عشوائية من 21 عضو هيئة تدريس بإحدى الجامعات والمطلوب اختبار فرض العدم H_0 : M = 25.1 ضد الفرض البديل $H_1 : M > 25.1$ باستخدام مستوى معنوية $\alpha = 0.05$.

جدول (١٢-٢٣)

11.1	12.9	14.0	19.4	19.9	22.2	23.6	24.1	25.2	26.7	25.1
28.1	29.6	30.7	32.2	33.9	34.8	38.9	40.1	50.5	50.6	

لحساب عدد الإشارات السالبة k نحدد إشارات الفروق بين مشاهدات العينة والوسيط حيــث وجدت كالآتي :

- - - - - - - - + + 0 + + + + + + + + + +

وحيث أن لدينا مشاهدة تساوى صفر وبعد استبعادها يصبح حجم العينة هو 20 مشـــاهدة . أى أن عدد الإشارات السالبة k = 8 . لمستوى معنوية $\alpha = 0.05$ نرفض H_0 إذا كانت :

$$P(K \le k \mid n, .5) \le \alpha .$$

نحسب :

$$P(K \le 8 \mid 20, 0.5) = \sum_{x=0}^{8} b(x;\ 20,\ 0.5) = 0.252.$$

حيث **0.252** تستخرج من جدول ذي الحدين في ملحق (١) عند **n = 20** و **p = 0.5** . بما أن **0.252** أكبر من **0.05** فإننا نقبل H_0 .

مثال (١٢-١٠) في مصنع للسكر تستخدم إحدى الماكينات لتعبئة السكر في أكياس بحيث أن وزن كل كيس 5 كيلو . أختيرت عينة عشوائية من الحجم **n = 10** من هذه الماكينة والأوزان معطاة في جدول (١٢-٢٤) .

المطلوب اختبار فرض العدم $H_0 : M = 5$ ضد الفرض البديل $H_0 : M_0 < 5$ وذلك عند مستوى معنوية **α= 0.05** .

جدول (١٢- ٢٤)

4.1	4.2	4.3	4.5	4.7	4.8	4.9	4.9	5.6	4.9

الحل . لحساب عدد الإشارات الموجبة نحدد إشارات الفروق بين مشاهدات العينة والوسيط حيث وجدت كالآتي :

- - - - - - - - + -

أي أن عدد الإشارات الموجبة **k = 1** . لمستوى معنوية **α= 0.05** نرفض H_0 إذا كانت :

$$P(K \le k \mid n, 0.5) \le \alpha .$$

نحسب :

$$P(K \le 1 \mid 10, 0.5) = \sum_{x=0}^{1} b(x;\ 10,\ 0.5) = 0.011.$$

حيث **.011** تستخرج من جدول ذي الحدين في ملحق (١) عند **n = 10 , p = .5** . وبما أن **0.011** اقل من **0.05** نرفض H_0 .

للعينات من الحجم 12 أو أكبر فإنه يمكننا استخدام التوزيع الطبيعي كتقريب لتوزيع ذي الحدين حيث :

$$z = \frac{(k \pm .5) - 0.5\ n}{0.5\sqrt{n}}$$

هي قيمة لمتغير عشوائي Z تقريباً يتبع التوزيع الطبيعي القياسي. سوف يستخدم $k + 0.5$ إذا كانت k أقل من $\frac{n}{2}$ و $k - 0.5$ إذا كانت k أكبر من $\frac{n}{2}$. في مثال (١٢-٩) فإن :

$$z = \frac{(k \pm 0.5) - 0.5\ n}{0.5\sqrt{n}}$$

$$= \frac{(8 + 0.5) - 0.5(20)}{0.5\sqrt{20}} \approx -0.67.$$

لمستوى معنوية $\alpha = 0.05$ فإن $z_{.05} = 1.645$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) . منطقة الرفض $Z < -1.645$. وبما أن z المحسوبة (0.67 -) تقع في منطقة القبول نقبل H_0 .

(٢-٧) اختبار الإشارة لعينتين مرتبطتين (عينة مزدوجة)

The Sign -Test for Two related Sample

يمكن تعديل اختبار الإشارة لعينة واحدة واستخدامه في حالة العينتين المرتبطتين عندما لا يتحقق الاعتدال . بفرض أن لدينا n من أزواج المشاهدات المستقلة $(x_i\ ,\ y_i)\ ; i = 1,2,...,n$ فإننا نستبدل كل زوج من المشاهدات بإشارة موجبة إذا كانت المشاهدة الأولى أكبر من المشاهدة الثانية وبإشارة سالبة إذا كانت المشاهدة الأولى أصغر من المشاهدة الثانية ونهمل الزوج الذي فيه المشاهدات متساويتان. لاختبار فرض العدم $H_0 : \mu_D = 0$ ضد الفرض البديل $H_1 : \mu_D \neq 0$ أو الفرض البديل $H_1 : \mu_D < 0$ أو الفرض البديل $H_1 : \mu_D > 0$ نتبع الخطوات المبينة في البند السابق .

مثال (١٢-١١) يعطى الجدول (١٢-٢٥) كمية مركب في دم عينة عشوائية من 21 حيوان قبل وبعد إعطائهم دواء لتدعيم النقص في دورة ما.

المطلوب اختبار فرض العدم $H_0 : \mu_D = 0$ ضد الفرض البديل $H_1 : \mu_D < 0$ وذلك عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢-٢٥)

الحيوان	قبل x_i	بعد y_i	إشارة $(x_i - y_i)$
1	2.5	2.6	-
2	1.2	1.5	-
3	2.9	2.9	0
4	3.1	2.0	+
5	3.1	2.3	+
6	1.1	1.5	-
7	1.5	1.6	-

←

8	4.1	3.1	+
9	2.1	1.4	+
10	2.4	2.5	-
11	1.3	1.4	-
12	2.8	2.9	-
13	3.5	2.4	+
14	3.6	2.1	+
15	1.1	1.3	-
16	1.6	1.7	-
17	4.2	3.2	+
18	2.2	1.5	+
19	2.5	2.1	+
20	1.3	1.1	+
21	1.3	1.5	-

الحل . إشارات الفروق بين أزواج المشاهدات $(x_i \;,\; y_i)$ $i = 1,2,\ldots,n$ معطاة في جدول (١٢-٢٥) وبعد إهمال الفرق المساوي للصفر وتعديل حجم العينة إلى 20 =n فإن عدد الإشارات الموجبة k = 10 . نرفض H_0 عند مستوى معنوية 0.05 =α إذا كانت :

$$P(K \le k \mid n, 0.5) \le \alpha .$$

نحسب : $P(K \le 10 \mid 20, 0.5)$

$$= \sum_{x=0}^{10} b(x;\; 20,\; 0.5) = 0.588.$$

حيث أن القيمة 0.588 تستخرج من جدول ذي الحدين في ملحق (١) عند n = 20 , p =0.5 . وبما أن 0.588 أكبر من 0.05 فإننا نقبل H_0 .

مثال (١٢-١٢) ترغب شركة في اختبار مدي اهتداء نوعين من إطارات السيارات A , B . ولهذا قامت بتركيب الإطارات من النوع B وبعد استخدامها نفس المسافة وقياس مدي الاهتراء حصلت الشركة على البيانات المعطاة في جدول (١٢-٢٦) والمطلوب اختبار فرض العدم $H_0 : \mu_D = 0$ ضد الفرض البديل $H_1 : \mu_D \neq 0$ عند مستوى معنوية α= 0.05 .

السيارة	1	2	3	4	5	6	7	8	9	10
A x_i	23	20	26	25	48	26	25	24	16	20
B y_i	20	30	16	33	23	24	8	21	13	18
إشارة $(x_i - y_i)$	+	-	+	-	+	+	+	+	+	+

الحل . من جدول (١٢-٢٦) فإن عدد الإشارات السالبة k=2. لمستوى معنوية 0.05 =α
نرفض H_0 إذا كان :

$$P(K \leq k \mid n, 0.5) \leq \frac{\alpha}{2}.$$

نحسب :

$$P(K \leq 2 \mid 10, 0.5)$$
$$= \sum_{x=0}^{2} b(x;\ 10,\ 0.5) = 0.055.$$

حيث 0.055 مستخرجة من جدول ذي الحدين في ملحق (١) عند $p = 0.5$, $n = 10$ وبما أن 0.055 أكبر من 0.025 . فإننا نقبل H_0 .

(١٢ – ٨) اختبار إشارة الرتب The Signed – ranks Test

يعتمد اختبار الإشارة لعينة واحدة والذي تناولناه في البند (١٢-٦) على الفرق بين قيم مشاهدات العينة والوسيط المفترض مع إغفال قيمة الفروق والذي يؤدى إلى أضعاف الاختبار . لذلك أقترح العالم Wilcoxon إختباراً لامعلمياً آخر أطلق عليه أسمه يعتمد على إشارة الفرق وقيمة الفرق حيث يعطى وزناً أكبر للإشارة التي تصاحب فرقاً كبيراً والعكس صحيح . يشترك هذا الاختبار مع اختبار الإشارة في أنه يمكن أن يستخدم كاختبار للمتوسط عندما يكون المجتمع موضع الدراسة متماثل .

تتكون البيانات اللازمة للتحليل من عينة عشوائية من الحجم n من المشاهدات المستقلة $x_1, x_2, \ldots, x_n$ والمختارة من مجتمع متصل. فرض العدم والفرض البديل سوف يكونان على الشكل :

$$H_0 : M = M_0,$$
$$H_1 : M \neq M_0.$$

لحساب قيمة الإحصاء الذي يعتمد عليه قرارنا ، عند مستوى معنوية α ، نتبع الخطوات التالية :

(أ) نحدد إشارة وقيمة الفروق $D_i = (x_i - M_0); i = 1,2,\ldots,n$.

(ب) إذا كان الفرق مساوياً للصفر تستبعد المشاهدة من التحليل ويعدل حجم العينة بطرح عدد يساوى عدد المشاهدات التي تساوي الوسيط .

(ج) نهمل إشارة الفروق مؤقتا ونرتب الفروق تصاعديا بمعنى آخر نعطى رتب للقيم $|D_i|$ أي القيم المطلقة للفروق وإذا كان هناك فروق متساوية ، أي تداخلات ties ، فإننا نعطيها متوسط الرتب التي كانت ستأخذها لو أنها كانت مختلفة .

(د) تعاد الإشارات إلى الرتب المناظرة ونوجد مجموع رتب الإشارات السالبة ونرمز له بالرمز t_- ونوجد مجموع رتب الإشارات الموجبة ونرمز له بالرمز t_+ ويمكن إيجاد كل منهما بدلالة الآخر من العلاقة :

$$t_+ = \frac{n(n+1)}{2} - t_-$$

نحسب مجموع الرتب السالبة أو مجموع الرتب الموجبة أيهما اقل ونرمز له بالرمز t'' أي أن:

$$t'' = \min(t_+, t_-).$$

والتي تمثل قيمة لإحصاء . يستخدم الجدول في ملحق (١١) لاستخراج القيم الحرجة لهذا الإحصاء لعينات من الحجم 3 وحتى الحجم 25 وذلك لاختبار من جانب أو جانبين . سنرمز للقيم الجدولية بالرمز $d(n,\alpha'')$, $d(n,\alpha')$ عندما يكون الاختبار من جانب أو جانبين على التوالي. للبديل $H_0 : M \neq M_0$ و لمستوى معنوية α نرفض H_0 إذا كان $t'' < d(n,\alpha'')$. للفرض البديل $H_0 : M > M_0$ اهتمامنا سوف يكون في مجموع الرتب السالبة t_-. لمستوى معنوية α نرفض H_0 إذا كان $t_- < d(n,\alpha')$. للفرض البديل $H_0: M < M_0$ اهتمامنا سوف يكون في مجموع الرتب الموجبة t_+. لمستوى معنوية α نرفض H_0 إذا كان $t_+ < d(n,\alpha')$.

مثال (١٢-١٣) اختيرت عينة عشوائية من 15 شخصا ممن يمتلكون منزلا في مدينة ما . وقد تم سؤالهم عن مقدار الزيادة في فاتورة الضرائب السنوية بالدولار. البيانات التي تم الحصول عليها معطاة في جدول (١٢- ٢٧) المطلوب اختبار فرض العدم $H_1 : M = 500$ ضد الفرض البديل $H_1 : M \neq 500$ عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢ - ٢٧)

مقدار الزيادة	$D_i = (x_i - 500)$	رتب $\|D_i\|$	$\|D_i\|$ مضروبا في إشارة D_i
494.4	-5.6	8	-8
510.8	+10.8	13	+13
487.5	-12.5	14	-14
493.2	-6.8	11	-11
502.6	+2.6	4	+4
500.0	0	يستبعد	–
495.9	-4.1	6	-6
498.2	-1.8	3	-3
501.6	+1.6	2	+2
497.3	-2.7	5	-5
492.0	-8.0	12	-12

504.3	+4.3	7	+7
499.2	-0.8	1	-1
493.5	-6.5	10	-10
505.8	+5.8	9	+9

الحل . من جدول (١٢-٢٧) ومن العمود الأخير (على اليمين) نحسب مجموع الرتب السالبة والموجبة كالتالي :

$$t_{-} = 8 + 14 + 11 + 6 + 3 + 5 + 12 + 1 + 10 = 70,$$
$$t_{+} = 13 + 4 + 2 + 7 + 9 = 35$$

وحيث أن هناك فرق مساوياً للصفر فإننا نهمله ونعدل حجم العينة تبعاً لذلك أي أن n=14 .

$$t'' = \min (t_{+} , t_{-}) = 35.$$

لمستوى معنوية $\alpha = 0.05$ فإن $d\ (14, .049) = 22$ حيث $\alpha'' = 0.049$ وعند n=14 من الجدول في ملحق (١١) وبما أن $t'' = 35$ أكبر من 22 نقبل H_0 .

عندما تكون n أكبر من 25 ، فإننا لا نستطيع استخدام الجدول في ملحق (١١) لتقدير المعنوية للقيمة المحسوبة للإحصاء.

للعينات الكبيرة فإن :

$$t^{*} = \frac{t'' - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}.$$

تمثل قيمة للإحصاء T^{*} الذي تقريبا يتبع التوزيع الطبيعي القياسي . للاختبارات من جانب واحد فإنه يمكن استبدال t'' في صيغة t^{*} بالقيمة t_{+} أو t_{-} حسب الفرض المستخدم. في حالة وجود تداخلات فلا بد من تصحيح قيمة الإحصاء t^{*} . بفرض أن u تمثل عدد الفروق المطلقة المتساوية لرتبة لا تساوي الصفر فإن معامل التصحيح سوف يكون كالتالي :

$$\frac{\Sigma u^{3} - \Sigma u}{48}.$$

حيث يطرح هذا المقدار من المقام تحت الجذر من صيغة t^{*} . وعلى ذلك يصبح المقام في صيغة t^{*} تحت الجذر هو :

$$\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\Sigma u^{3} - \Sigma u}{48}}.$$

سوف نوضح كيفية تصحيح التداخلات باستخدام البيانات في الجدول (١٢ – ٢٨) .

جدول (١٢–٢٨)

$\lvert D_i \rvert$	2	2	5	6	6	6	7	7	8	8	8	8
الرتب	1.5	1.5	3	5	5	5	7.5	7.5	10.5	10.5	10.5	10.5

من جدول (١٢– ٢٨) يمكن الحصول على الجدول (١٢ – ٢٩) حيث u هو عدد القيم في كل فئة بها تداخلات .

جدول (١٢– ٢٩)

U	2	3	2	4
U^3	8	27	8	256

$$\Sigma u^3 = 299 \quad , \quad \Sigma u = 11$$

وعلى ذلك معامل التصحيح للتداخلات هو :

$$\frac{\Sigma u^3 - \Sigma u}{48} = \frac{299 - 11}{48} = 6.$$

(١٢– ٩) اختبار Mann – Whitney – Wilcoxon

يشترط في اختبار t الذي يخص الفرق بين متوسطي مجتمعين ، الذي تناولناه في الفصل التاسع ، أن المجتمعين اللذين اخترنا منهما العينتين يتبعان توزيعاً طبيعياً . عندما لا يتوفر هذا الشرط فإن اختبار **Mann-Whitney** يكون البديل .

تتكون البيانات اللازمة للتحليل من عينة عشوائية من الحجم n_1 من المشاهدات $x_1, x_2, ..., x_n$ من المجتمع الأول المتصل . أيضا نختار عينة عشوائية أخرى من الحجم n_2 من المشاهدات $y_1, y_2, ..., y_{n_2}$ من المجتمع الثاني المتصل ويشترط أن تكون العينتين مستقلتين. فرض العدم والفرض البديل سوف يكونان على الشكل :

H_0 : المجتمعين لهما نفس التوزيع .

H_1 : قيم x's تتجه لأن تكون أصغر من قيم y's .

لإجراء الاختبار نقوم بدمج مشاهدات العينتين معا في عينة واحدة ثم نقوم بترتيب المشاهدات تصاعدياً فنعطي الرتبة 1 لأصغر مشاهدة والرتبة 2 للمشاهدة التي تليها في العينة وهكذا حتى المشاهدة الأخيرة والتي تمثل أكبر قيمة حيث تعطى الرتبة $n_1 + n_2$. إذا ظهرت مشاهدات متساوية في العينة (تداخلات) فإننا نرتب المشاهدات كما لو أنها ليست فيها مشاهدات متساوية في العينة ثم نحسب الوسط الحسابي لرتب المشاهدات في فئة المشاهدات المتساوية في القيمة ونعتبر الوسط الحسابي رتبة لكل مشاهدة في الفئة . نحسب القيمة :

$$w = s - \frac{n_1(n_1+1)}{2}$$

حيث s تمثل مجموع الرتب للعينة المختارة من المجتمع الأول و w قيمة للإحصاء W وذلك تحت فرض أن H_0 صحيح . لمستوى معنوية α فإن منطقة الرفض $W < w_\alpha$ حيث w_α هي القيمة الحرجة للإحصاء W وتستخرج من الجدول في ملحق (١٢) عند n_1, n_2 ومستويات معنوية مختلفة. للفرض البديل H_0 : قيم x's تتجه لأن تكون أكبر من قيم y's فإن منطقة الرفض $W > w_{1-\alpha}$ حيث $w_{1-\alpha}$ تحسب كالتالي :

$$w_{1-\alpha} = n_1 n_2 - w_\alpha$$

للفرض البديل H_1 : المجتمعان يختلفان بالنسبة للموقع ، فإن منطقة الرفض $W > w_{1-\frac{\alpha}{2}}$ أو $W < w_{\frac{\alpha}{2}}$ حيث :

$$w_{1-\frac{\alpha}{2}} = n_1 n_2 - w_{\frac{\alpha}{2}}$$

مثال (١٢-١٤) يعطى الجدول (١٢-٣٠) أزمنة الفشل لنوعين من الأجهزة الإلكترونية والمطلوب اختبار فرض العدم H_0 : المجتمعين لهما نفس التوزيع ضد الفرض البديل H_1 : المجتمعين ليس لهما نفس التوزيع (عند مستوى معنوية $\alpha = 0.05$) .

جدول (١٢-٣٠)

x	23	261	87	7	120	14	62	47	225	71	246	21
y	55	320	56	104	220	239	47	246	176	182	33	

من الجدول (١٢-٣١) يتم حساب مجموع الرتب للعينة الأولي وهي s=124 .

جدول (١٢ - ٣١)

	x 7	x 14	x 21	x 23	y 33	y 47	x 47	y 55	y 56	x 62	x 71	x 87
الرتب	1	2	3	4	5	6.5	6.5	8	9	10	11	12
	y 104	x 120	y 176	y 182	y 220	x 225	y 239	x 246	y 246	x 261	y 320	
الرتب	13	14	15	16	17	18	19	20.5	20.5	22	23	

وعلى ذلك فإن :

$$w = s - \frac{n_1(n_1+1)}{2}$$
$$= 124 - \frac{12(12+1)}{2}$$
$$= 124 - 78 = 46.$$

من الجدول في ملحق (١٢) فإن $w_{.025}=34$ عند $n_1=12$ و $n_2=11$ وبما أن :

$$w_{1-\frac{\alpha}{2}} = n_1 n_2 - w_{\frac{\alpha}{2}}$$

فإن :

$$w_{0.975} = (12)\,(11) - 34 = 98.$$

منطقة الرفض $W > 98$ أو $W < 34$ وبما أن $w = 46$ تقع في منطقة القبول نقبل H_0 .

عندما n_2 , n_1 أكبر من 20 فإننا لا نستطيع استخدام الجدول في ملحق (١٢) وفي هذه الحالة يمكن حساب القيمة :

$$z = \frac{w - \frac{n_1 n_2}{2}}{\sqrt{(n_1 n_2)(n_1 + n_2 + 1)/12}}$$

والتي تمثل قيمة للمتغير العشوائي Z وهو تقريباً يتبع التوزيع الطبيعي القياسي وذلك تحت فرض أن H_0 صحيح .

بالنسبة لمشكلة التداخلات والتي قد تحدث داخل كل مجموعة أو بين قيم المجموعتين فقد تم إثبات أن التداخلات داخل المجموعة ليس لها تأثير على قيمة الإحصاء ولكن وجود تداخلات بين المجموعتين يؤثر على النتائج . في هذه الحالة لا بد من عمل تصحيح لقيمة z . بفرض أن u ترمز لعدد التداخلات لرتب معطاة فإن معامل التصحيح للتداخلات يحسب من الصيغة التالية :

$$\frac{n_1 n_2 (\Sigma u^3 - \Sigma u)}{12(n_1 + n_2)(n_1 + n_2 - 1)}$$

والتي تطرح من المقام في صيغة z تحت الجذر . وعلى ذلك المقام في صيغة z يصبح :

$$\sqrt{\frac{(n_1 n_2)(n_1 + n_2 + 1)}{12} - \frac{n_1 n_2 (\Sigma u^3 - \Sigma u)}{12(n_1 + n_2)(n_1 + n_2 - 1)}}.$$

(١٢-١٠) اختبار Kruskal – Wallis

يعتبر اختبار Kruskal – Wallis من أكثر الاختبارات اللامعلمية شيوعاً لإجراء تحليل التباين في حالة التصنيف الأحادي. تتكون البيانات اللازمة للتحليل من k من العينات العشوائية من الحجم $n_1, n_2, \ldots, n_k$ على أن تكون المشاهدات مستقلة سواء بين أو داخل المعالجات كما أن المجتمعات التي اختيرت منها العينات تكون من النوع المتصل . فرض العدم والفرض البديل سوف يكونان على الشكل :

H_0 : التوزيعات للمجتمعات التي عددها k متطابقة .

H_1 : التوزيعات للمجتمعات التي عددها k ليس لها نفس الوسيط .

في هذا الاختبار نقوم بدمج مشاهدات العينات في عينة واحدة وإعطاء رتب لهذه المشاهدات حيث R_i ترمز إلى مجموع الرتب للمشاهدات التي تنتمي إلى المجموعة رقم i والتي عدد مشاهداتها n_i حيث أن $N = \sum_{i=1}^{k} n_i$ تمثل العدد الكلي للمشاهدات الناتجة من k من العينات . قيمة الإحصاء الذي يعتمد عليه قرارنا هو :

$$h = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{1}{n_i} \left[R_i - \frac{n_i(N+1)}{2} \right]^2$$

والتي يمكن تبسيطه بالصيغة التالية :

$$h = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1).$$

حيث أن h هو قيمة للإحصاء H بافتراض أن H_0 صحيح .تستخرج القيم الحرجة للإحصاء H من الجدول في ملحق (١٣) عند مستويات مختلفة من المعنوية بشرط أن $n_i \le 5, i = 1,2,3$. إذا كان عدد العينات أو عدد المشاهدات داخل كل عينة غير متوفر في الجدول فقد وجد بالبرهان أن H تقريبا تتبع توزيع χ^2 تحت شرط أن عدد المشاهدات في كل عينة لا تقل عن 5 أي أن $n_i \ge 5, i = 1,2,...,k$. لمستوى معنوية α فإن منطقة الرفض $\chi^2 > \chi^2_\alpha$ حيث أن χ^2_α تستخرج من جدول توزيع χ^2 عند مستوى معنوية α ودرجات حرية k-1 . نرفض H_0 إذا وقعت h المحسوبة في منطقة الرفض .

في حالة وجود تداخلات لا بد من تصحيح الإحصاء H إلى الإحصاء H' حيث :

$$H' = \frac{H}{C},$$

$$C = 1 - \frac{\Sigma(u_i^3 - u_i)}{N^3 - N}$$

حيث u_i هو عدد القيم في كل فئة بها قيم متساوية و N هي القيمة الناتجة من دمج العينات التي عددها k أي أن $N = \sum_{i=1}^{k} n_i$. لمستوى معنوية α فإن منطقة الرفض $H' > \chi^2_\alpha$ حيث:

χ^2_α تستخرج من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية k-1 .إذا وقعت h' المحسوبة $(h' = \frac{h}{C})$ في منطقة الرفض نرفض H_0. إذا كانت القيم المتداخلة قليلة في العينات فإن H' تقترب من **H** .

مثال (١٢-١٥) استخدمت ثلاثة طرق تعليمية مختلفة لتعليم ثلاثة مجموعات متشابهة من الطلبة وكانت درجات الامتحان النهائي معطاة في جدول (١٢-٣٢)

جدول (١٢- ٣٢)

المجموعة A	20	37	39	41	45
المجموعة B	43	46	48	53	
المجموعة C	31	38	44		

هل توجد فروق معنوية بين الطرق الثلاثة ؟وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل .

H_0 : توزيعات المجتمعات الثلاثة التي اختيرت فيها العينات الثلاثة متطابقة .

H_1 : توزيعات المجتمعات الثلاثة ليس لها نفس الوسيط.

للحصول على القيمة **h** تم دمج الثلاث عينات معا ووضع رتب للعينة المشتركة وتحديد رتب كل عينة والنتائج معطاة في الجدول (١٢-٣٣).

جدول (١٢-٣٣)

	الرتب	R_i
A	1 3 5 6 9	24
B	7 10 11 12	40
C	2 4 8	14

وعلى ذلك فإن قيمة **h** هي :

$$h = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i}{n_i} - 3(N+1)$$

$$= \frac{12}{N(N+1)} \left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} \right] - 3(N+1)$$

$$= \frac{1}{13} \left[\frac{24^2}{5} + \frac{40^2}{4} + \frac{14^2}{3} \right] - 3(13)$$

$$= \left(\frac{1}{13}.580.53\right) - 39$$
$$= 44.66 - 39 = 5.66.$$

لمستوى معنوية $\alpha = 0.05$ القيمة الحرجة للإحصاء H والمستخرجة من الجدول في ملحق (١٣) هي 5.6308 عند $n_1 = 5, n_2 = 4, n_3 = 3$. منطقة الرفض $H > 5.6308$

وبما أن $h = 5.66$ تقع في منطقة الرفض نرفض H_0 .

مثال (١٢-١٦) في دراسة على نوع معين من الطحالب في بحيرة صغيرة قام باحث بمقارنة النسبة المئوية لعدد الطحالب من نوع ما في منطقة محددة خلال فصل الربيع ومنتصف الصيف وآخر الخريف والبيانات معطاة في جدول (١٢-٣٤) .

جدول (١٢ -٣٤)

الربيع %	9.6	11.2	11.6	11.7	12.8	12.9	15.8	22.7	24.6	32.5
منتصف الصيف %	4.8	7.6	7.6	9.2	9.6	21.1	24.6	25.6	26.4	32.8
آخر الخريف %	5.4	6.5	7.1	8.0	8.8	9.5	10.2	10.7	11.3	11.7

النتائج اللازمة لحساب h في جدول (١٢ -٣٥)

جدول (١٢ -٣٥)

											R_i
الربيع	11.5	15	17	18.5	20	21	22	24	25.5	29	203.5
منتصف الصيف	1	5.5	5.5	9	11.5	23	25.5	27	28	30	166
آخر الخريف	2	3	4	7	8	10	13	14	16	18.5	95.5

$$h = \frac{12}{N(N+1)}\left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3}\right] - 3(N+1)$$

$$= \frac{12}{30(31)}\left[\frac{203.5^2}{10} + \frac{166^2}{10} + \frac{95.5^2}{10}\right] - 3(31)$$
$$= 7.759.$$

ولوجود تداخلات لابد من تحويل h إلي h' . نحسب $u^3 - u$ لكل تداخل ثم نحسب $\Sigma(u^3-u)$ كالتالي :

$$(8-2) + (8-2) + (8-2) + (8-2) = 24.$$

معامل التصحيح يحسب من الصيغة التالية :

$$1 - \frac{\Sigma(u^3 - u)}{N^3 - N}$$

$$= 1 - \frac{24}{27000 - 30} = 0.9991.$$

وعلى ذلك :

$$h' = \frac{7.759}{0.9991} = 7.76599.$$

$\chi^2_{.05} = 5.992$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) عند درجات حرية $k-1=3-1=2$. منطقة الرفض $X^2 > 5.992$. وبما أن $h' = 7.76599$ تقع في منطقة الرفض نرفض H_0 .

(١٢-١١) اختبار الدورات Test Runs

يفيد هذا الاختبار في مجالات كثيرة منها المشاكل البيولوجية . فقد يرغب باحث ما في معرفة ما إذا كان عدد حالات الإصابة بمرض الملاريا ، مثلاً ، يتغير عشوائياً من سنة إلي سنة أخرى أم أن هناك عوامل غير عشوائية تؤدى إلى نقص أو زيادة عدد حالات الإصابة بهذا المرض. لإجراء الاختبار نفترض أن لدينا متتابعة من المشاهدات المسجلة تبعاً لترتيب حدوثها وأن المشاهدات يمكن تقسيمها إلي نوعين (ليكن a , b) . يعتمد هذا الاختبار على متغير يطلق عليه أسم الدورة run حيث تعرف بأنها مجموعة الأحداث المتشابهة التي يسبقها أو يتبعها نوع آخر مخالفا من الأحداث أو لا يتبعها أو لا يسبقها أية أحداث. عدد الإحداث في الدورة يطلق عليها طول الدورة (يمكن أن تحتوى الدورة على حدث واحد). بفرض أن لدينا البيانات التالية والتي تم توليدها على الحاسب الآلي .

0.1, 0.4, 0.2, 0.8, 0.6, 0.9, 0.3, 0.4, .01, 0.2

بفرض أن a تمثل الرقم الذي أقل من 0.5 و b تمثل الرقم الذي أكبر من 0.5 والتي تعطى المتتابعة :

a a a   b b b   a a a a

في هذه الحالة لدينا ثلاث دورات . سوف نرمز لعدد الدورات بالرمز r' ، أي أن $r' = 3$.
أيضا n_1 سوف ترمز لعدد قيم a و n_2 سوف ترمز لعدد قيم b . فرض العدم والفرض البديل سوف يكونان على الشكل

H_0 : النوعان يقعان بعشوائية .

H_1 : النوعان لا يقعان بعشوائية .

بفرض أن H_0 صحيح فإن r' هي قيمة للإحصاء R' . القيم الحرجة السفلي للإحصاء R' تستخرج من الجدول في ملحق (١٤) والقيم الحرجة العليا للإحصاء R' تستخرج من الجدول في ملحق (١٥) وذلك عند مستوى معنوية $n_1, n_2, \alpha = 0.05$. لتكن r_1 القيمة السفلي و r_2 القيمة العليا عند $n_1, n_2, \alpha = 0.05$ ، منطقة الرفض سوف تكون $R' > r_2$ أو $R' < r_1$. إذا وقعت القيمة r في منطقة الرفس نرفض H_0 . للفرض البديل H_1 : النوعان لا يقعان بعشوائية لوجود عدد كبير من الدورات ولمستوى معنوية $\frac{\alpha}{2}$ فإن منطقة الرفض $R' > r_2$.
للفرض البديل H_1 : النوعان لا يقعان بعشوائية لوجود عدد قليل من الدورات ولمستوى معنوية $\frac{\alpha}{2}$ فإن منطقة الرفض $R' < r_1$. إذا كانت n_1 , n_2 أكبر من 20 (كليهما أو أيهما) فإن الجدولان في ملحق (١٤) و (١٥) لا يصلحان للاستخدام و لقد وجد بالبرهان أن :

$$z = \frac{r' - \left[\frac{2n_1n_2}{n_1+n_2} + 1\right]}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}}.$$

قيمة لمتغير عشوائي Z تقريباً يتبع التوزيع الطبيعي القياسي .

مثال (١٢- ١٧) إذا كان عدد حالات الملاريا المسجلة في منطقة ما خلال عدة سنوات مسجلة في الجدول (١٢-٣٦) .

جدول (١٢-٣٦)

51	82	64	32	11	12	54	71	90	101	84	72	45	20	74	15

هل عدد حالات الإصابة تأخذ شكل عشوائي أم لا ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل . H_0: البيانات عشوائية .

H_1 : البيانات غير عشوائية .

عادة يعتبر الوسيط هو القيمة التي تستخدم لتقسيم البيانات في العينة إلي نوعين a وb . للبيانات في جدول (١٢–٣٦) فإن الوسيط 59 . سوف نرمز للقيمة التي أصغر من قيمة الوسيط بالرمز a والقيمة التي أكبر من قيمة الوسيط بالرمز b وبذلك نحصل على المتابعة التالية :

a bb aaaa bbbbb aa b a

وعلى ذلك :

$$r' = 7 \;,\; n_2 = 8 \;,\; n_1 = 8$$

لمستوى معنوية $\alpha = 0.05$ فإن $r_1 = 4$ والمستخرجة من الجدول في ملحق (١٤) عند $n_1 = 8$ $n_2 = 8$.و$r_2 = 14$ والمستخرجة من الجدول في ملحق (١٥) عند $n_2 = 8$ $n_1 = 8$. منطقة الرفض $R' > 14$ أو $R' < 4$. وبما أن $r' = 7$ فإنها تقع في منطقة القبول وبذلك نقبل فرض العدم .

مثال (١٢ – ١٨) إذا كان لدينا 34 شخصا وبفرض أن M ترمز للذكر و F ترمز للأنثى وكانت النتائج كالآتي :

FF MMMMMMMMM FF M FF MMMMMMMMM FFFFF MMM F

المطلوب تحديد هل العينة عشوائية أم لا ؟ وذلك عند مستوى معنوية $\alpha = 0.05$:

الحل. H_0 : العينة عشوائية.

H_1 : العينة غير عشوائية.

من بيانات العينة نجد أن :

$r' = 9$, $n_2 = 22$ (عدد M) و $n_1 = 12$ (عدد F) وبما أن $n > 20$ فإننا نستخدم التقريب الطبيعي وعلى ذلك :

$$z = \frac{r' - \left[\dfrac{2n_1n_2}{n_1+n_2} + 1\right]}{\dfrac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$$

$$= \frac{9 - \left[\dfrac{2(12)(22)}{12+22} + 1\right]}{\dfrac{2(12)(22)[2(12)(22) - 12 - 22]}{(12+22)^2(12+22-1)}}$$

$$= \frac{9-16.529}{6.8373702}$$
$$= -1.101.$$

لمستوى معنوية $\alpha = 0.05$ و $z_{.025} = 1.96$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) . منطقة الرفض $Z > 1.96$ أو $Z < -1.96$ وبما أن $z = -1.101$ تقع في منطقة القبول نقبل H_0 .

(١٢–١٢) معامل ارتباط سبيرمان للرتب

The Spearman Rank Correlation Coefficient

تناولنا في البند (١٠–٤) من الفصل العاشر اختبارات الفروض التي تخص معامل ارتباط المجتمع ρ تحت فرض أن X , Y متغيرين عشوائيين لهما توزيع طبيعي ثنائي . في حالة عدم تحقق الشرط السابق فإنه يمكننا استخدام معامل سبيرمان كإحصاء لاختبار عدم وجود علاقة (ارتباط) بين المتغيرين X , Y. أيضا يمكننا استخدام معامل سبيرمان كمقياس وصفي لقوة الارتباط بين متغيرين X , Y عندما تكون البيانات في العينة غير متوفرة في شكل بيانات رقمية ولكن يمكن تعيين رتب لها. لإجراء الاختبار نتبع الآتي :

(أ) تختار عينة عشوائية من الحجم n من أزواج المشاهدات الرقمية أو الوصفية . كل زوج من المشاهدات يمثل قراءتين مأخوذتين على نفس المفردة والمسماة وحدة الاقتران **unit of association** . أيضا قد تمثل البيانات مشاهدات مأخوذة من مجتمع ثنائي . سوف نرمز لأزواج المشاهدات كالتالي $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$.

(ب) نرتب قيم المشاهدات في العينة والتابعة للمتغير X تصاعديا وتعطي رتبة لكل قيمة مشاهدة بالنسبة لكل قيم المشاهدات الأخرى. سوف نرمز لرتبة المشاهدة رقم **i** ، x_i ، بالرمز $r(x_i)$. عندما $r(x_i)=1$ فهذا يعنى أن x_i تمثل أقل قيمة مشاهدة من قيم المتغير X في العينة .

(ج) نرتب قيم المشاهدات في العينة والتابعة للمتغير Y تصاعدياً وتعطى رتبة لكل قيمة مشاهدة بالنسبة لكل قيم المشاهدات الأخرى .سوف نرمز لرتبة المشاهدة رقم **j** ، y_i ، بالرمز $r(y_i)$.عندما $r(y_i) = 1$ فهذا يعنى أن y_i تمثل أقل قيمة مشاهدة من قيم المتغير Y في العينة .

(ح) عند حدوث تداخلات نعطى متوسط الرتب المتتالية بدلاً من الرتبة كالمعتاد .

(خ) إذا كانت البيانات وصفية بإمكاننا تحويلها إلي رتب .

قيمة الإحصاء الذي يعتمد عليه قرارنا هو معامل ارتباط سبيرمان والذي يحسب من الصيغة التالية :

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}$$

حيث:

$$\Sigma\ d_i^2 = \Sigma\ \left[r(x_i) - r(y_i)\right]^2.$$

لكل زوج من المشاهدات وعندما تكون رتبة x نفس رتبة y (ارتباط تام طردي) ، فإن كل الفروق d_i سوف تساوى صفر وعلى ذلك $r_s = +1$. إذا كانت رتبة كل متغير داخل كل زوج من المشاهدات عكس الآخر (ارتباط تام عكسي) ، أي إذا كان :

$$[\ r(x) = 1,\ \ r(y) = n\ ]\ ,\ [\ r(x) = 2\ ,\ \ r(y) = n-1\ ]\ ,\ \ldots,\ [\ r(x) = n\ ,\ r(y) = 1]$$

وذلك لأزواج المشاهدات التي عددها n فإن $r_s = -1$.على سبيل المثال إذا كان لدينا أزواج المشاهدات التالية :

$$(x_i, y_i) : (12,5),(11,6),(10,7),(9,8)$$

فإن الرتب تصبح :

$$r(x_i) : 4 \quad 3 \quad 2 \quad 1$$
$$r(y_i) : 1 \quad 2 \quad 3 \quad 4$$

وعلى ذلك $\Sigma\ d_i^2$ سوف تكون :

$$(3)^2 + (1)^2 + (-1)^2 + (-3)^2 = 20\ .$$

وبالتعويض في معادلة سبيرمان فإن :

$$r_s = 1 - [(6)\ (20)\ /\ (4)\ (15)\ ]$$
$$= 1 - 2 = -1\ \ .$$

معامل ارتباط سبيرمان لا يمكن أن يزيد عن 1 + ولا يمكن أن يقل عن 1 - . فرض العدم والفرض البديل سوف يكونان على الشكل :

H_0 : المتغيرين مستقلين .

H_1 : توجد علاقة بين المتغيرين في نفس الاتجاه أو الاتجاه المعاكس .

بفرض أن H_0 صحيح فإن r_s تمثل قيمة للإحصاء R_s الذي له توزيع احتمالي . القيم الحرجة $r^*_{s,\alpha}$ للإحصاء R_s تستخرج من الجدول في ملحق (١٦) لعينات من الحجم 4 وحتى الحجم 30 عن مستويات مختلفة من المعنوية . لمستوى معنوية α فإن منطقة الرفض $R_s > r^*_{s,\alpha/2}$ أو

: H_1 للفرض البديل H_0 . نرفض $R_s < -r^*_{s,\alpha/2}$. إذا وقعت r_s في منطقة الرفض فإننا نرفض
توجد علاقة بين المتغيرين في نفس الاتجاه فإن منطقة الرفض $R_s > r^*_{s,\alpha}$ وذلك عند مستوى معنوية α . للفرض البديل H_1 : توجد علاقة بين المتغيرين في اتجاه معاكس فإن منطقة الرفض $R_s < -r^*_{s,\alpha}$ وذلك عند مستوى معنوية α .

القرارات السابقة تستخدم عندما لا يكون هناك تداخل أو أن يكون عددها صغيراً . عندما يكون هناك تداخل و إذا كان عددها كبيراً (العدد الصغير للتداخلات لا يؤثر على r_s) فيجب إجراء تصحيح على r_s ونحتاج جداول خاصة لإجراء الاختبار سوف لا نتعرض لها . عندما يكون حجم العينة كبيراً (أكبر من 30) فإننا لا نستطيع استخدام الجداول ولكن تم إثبات أن :

$$z = r_s / \sqrt{n-1}$$

قيمة للمتغير العشوائي Z والذي تقريباً يتبع التوزيع الطبيعي القياسي وذلك بافتراض أن H_0 صحيح .

مثال (١٢ – ١٩) لدراسة العلاقة بين الهيموجلوبين X (مقاساً mg/100 ml) وعدد كرات الدم الحمراء Y بالمليون لكل ملليمتر مكعب ، اختيرت عينة عشوائية من 12 ذكر بالغ من مجتمع ما وتم قياس تركيزات الهيموجلوبين وعدد كرات الدم الحمراء لكل مفردة والبيانات معطاة في جدول (١٢ – ٣٧) .

جدول (١٢ –٣٧)

الشخص	الهيموجلوبين		كرات الدم الحمراء		d	d^2
	x	رتب x	y	رتب y		
1	15.2	7.5	5.1	9	-1.5	2.25
2	16.4	12	5.4	11	1	1
3	14.2	2	4.5	4	-2	4
4	13.0	1	4.2	1	0	0
5	14.5	3	4.3	2.5	0.5	0.25
6	16.1	11	6.1	12	-1	1
7	15.2	7.5	5.2	10	-2.5	6.25
8	14.8	5	4.3	2.5	2.5	6.25
9	15.7	10	4.7	6	4	16
10	14.9	6	4.8	7.5	-1.5	2.25
11	15.6	9	4.6	5	4	16
12	14.7	4	4.8	7.5	-3.5	12.25

المطلوب اختبار فرض العدم H_0 : المتغيرين مستقلين ضد الفرض البديل H_1 : توجد علاقة بين المتغيرين في نفس الاتجاه أو الاتجاه المعاكس وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل . $\Sigma d_i^2 = 67.5$ وعلى ذلك فإن :

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)}$$
$$= 1 - \frac{6(67.5)}{12(144 - 1)}$$
$$= 1 - 0.2360139 = 0.763986.$$

$r_s^* = 0.5804$ والمستخرجة من الجدول في ملحق (١٦) عند مستوى معنوية $\frac{\alpha}{2} = 0.025$, $n = 12$ منطقة الرفض $R_s > 0.5804$ أو $R_s < -0.5804$. وبما أن $r_s = 0.763986$ تقع في منطقة الرفض نرفض H_0.

مثال (١٢ – ٢٠) يعطى الجدول (١٢ – ٣٨) تقديرات 10 طلاب في كل من الإحصاء والرياضيات .

جدول (١٢ – ٣٨)

تقديرات الإحصاء	جيد جدا	جيد	مقبول	جيد	جيد جدا	ممتاز	جيد	ممتاز	جيد	جيد
تقديرات الرياضيات	جيد	جيد	مقبول	جيد جداً	جيد	جيد جدا	مقبول	ممتاز	جيد جدا	جيد جدا

أختبر فرض العدم H_0 : المتغيرين مستقلين .

ضد الفرض البديل :

H_1 : توجد علاقة بين المتغيرين في نفس الاتجاه .

وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل . من جدول (١٢ – ٣٨) يمكن الحصول على جدول (١٢ – ٣٩) .

جدول (١٢ – ٣٩)

رتب x	7.5	4	1	4	7.5	9.5	4	9.5	4	4	المجموع
رتب y	4	4	1.5	7.5	4	7.5	1.5	10	7.5	7.5	
D_i	3.5	0	-0.5	-3.5	3.5	2	2.5	-0.5	-3.5	-3.5	0
d_i^2	12.25	0	0.25	12.25	12.25	4	6.25	0.25	12.25	12.25	72

وعلى ذلك :

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}$$

$$= 1 - \frac{6(72)}{10(100-1)}$$

$$= 1 - 0.4363636 = 0.5636363.$$

$r^*_{s,0.05} = 0.5515$ والمستخرجة من الجـــدول في ملحـــق (١٦) عنـــد مســـتوى معنويـــة $n = 10, \alpha = 0.05$. منطقة الرفض $R_s > 0.5515$. وبمـــا أن $r_s = 0.5636363$ تقـــع في منطقة الرفض نرفض H_0 .

تماريـــــــن :

-١- اختير 178 رقم من أحد الجداول عشوائياً وكان التوزيع التكراري لهذه الأرقام كمـــا في الجدول التالي :

الرقم	1	2	3	4	5	6	7	8	9
التكرار المشاهد	17	18	22	20	15	29	22	20	15

أختبر فرض العدم ، عند مستوى معنوية $\alpha = 0.05$ ، أن البيانات الســـابقة يمكـــن توفيقـــها بالتوزيع المنتظم .

-٢- ألقيت عمله 100 مرة وتم الحصول على 63 صورة و 37 كتابة . هل هذه النتائج تتفـــق مع الفرض القائل أن العملة متزنة ؟ استخدم مستوى معنوية $\alpha = 0.05$.

-٣- طبقا لنظرية في علم الوراثة فإن تهجين الورود الحمراء والورود الصفراء ينتج ورود حمـــراء بنسبة 50% وورود برتقالية بنسبة 25% وورود صفراء بنسبة 25% . لاختبار صحة هـــذه النظرية أجريت 164 عملية تهجين فتم الحصول على 29 وردة حمراء و 90 وردة برتقاليـــة و 45 وردة صفراء هلي تدل هذه البيانات على صحة النظرية . استخدم مستوى معنوية $\alpha = 0.05$.

-٤-ألقيت زهرة 120 مرة متتالية وتم الحصول على النتائج التالية :

عدد النقاط	1	2	3	4	5	6
التكرار المشاهد	18	21	17	22	23	19

هل يمكن القول بأنه هذه الزهرة غير متحيزة باستخدام مستوى معنوية $\alpha = 0.01$.

-٥- ألقيت أربع قطع نقود عشرين مرة وتم الحصول على النتائج التالية :

عدد الصور	0	1	2	3	4
التكرار المشاهد	3	4	5	5	3

المطلوب استخدام اختبار مربع كاي لجودة التوفيق لاختبار ما إذا كانت النتائج تتفق مع توزيع ذي الحدين باحتمال نجاح ½ وذلك عند مستوى معنوية $\alpha = 0.01$.

-٦- اختيرت ثلاثة كرات من إناء يحتوى على 5 كرات حمراء و 3 كرات خضراء وتم تسجيل العدد x الذي يمثل عدد الكرات الحمراء ثم أعيدت الكرات في الإناء وكررت التجربة 112 مرة . النتائج التي تم الحصول عليها في الجدول التالي :

x	0	1	2	3
التكرار المشاهد	1	31	55	25

أختبر فرض العدم ، عند مستوى معنوية $\alpha = 0.01$ ، أن البيانات التي تم الحصول عليها من التجربة يمكن توفيقها بالتوزيع الهندسي الزائدي $h(x\ ;\ 8,3,5)$ حيث $x=0,1,2,3$.

-٧- إذا كانت درجات مادة الإحصاء لفصل دراسي خاص كالآتي :

الدرجة	A	B	C	D	E
التكرار المشاهد	14	18	32	20	16

أختبر فرض العدم ، عند مستوى معنوية $\alpha = 0.01$ ، أن البيانات في الجدول السابق يمكن توفيقها بتوزيع منتظم .

-٨- ألقيت عملة حتى ظهر وجه فإذا كان X يمثل عدد المحاولات حتى ظهور وجه . بعد تكرار التجربة 256 مرة تم الحصول على البيانات التالية :

x	1	2	3	4	5	6	7	8
التكرار المشاهد	136	60	34	12	9	1	3	1

أختبر فرض العدم ، عند مستوى معنوية $\alpha = 0.01$ ، أن البيانات في الجدول السابق يمكن توفيقها بالتوزيع الهندسي $g(x\ ,\ ½)\ ,\ x=1,2,3,...$.

-٩- قامت شركة للتأمين على السيارات بتسجيل البيانات الخاصة بعدد الحوادث x التي تعرضت لها السيارات المؤمن عليها والبيانات في الجدول التالي :

التكرار المشاهد	10	40	100	150	200	125	75	50	30	20

هل يمكن القول أن عدد الحوادث التي تتعرض لها السيارات المؤمن عليها تتبع توزيع بواسون وذلك عند مستوى معنوية $\alpha = 0.05$.

-١٠- البيانات في الجدول التالي تعطي التوزيع التكراري لمتغير X .

(أ) وفق منحنى التوزيع الطبيعي لهذه البيانات .

(ب) اختبر جودة التوفيق وذلك عند مستوى معنوية $\alpha = 0.05$.

حدود الفئة	0-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16
التكرار	15	20	35	25	35	40	20	15

-١١- البيانات التالية تعطى عدد الوفيات بسبب تعاطى جرعات زائدة من مشروب كحلي وذلك عند أعمار مختلفة :

العمر	9-15	16-22	23-29	30-36	37-43	44-50	51-57
التكرار	40	35	32	10	13	13	4

المطلوب (أ) وفق منحنى التوزيع الطبيعي لهذه البيانات؟

(ب) اختبر جودة التوفيق عند مستوى معنوية $\alpha = 0.01$

-١٢- في معركة حربية كان عدد المناطق m_y التي تستقبل y ضربات كالتالي :

y	0	1	2	3	4	≥ 5
m_y	229	211	93	35	7	1

اختبر فرض العدم أن عدد الضربات متغير عشوائي يتبع توزيع بواسون بمعلمة μ عند مستوى معنوية $\alpha = 0.05$ افترض أن $\bar{x} \approx 3.25$ عند مستوى معنوية $\alpha = 0.05$.

-١٣- إذا كانت أزمنة الحياة لـ 100 خلية ضوئية flashlight ، تم ملاحظتها موزعة كالتالي:

عدد الدقائق	0-706	706-746	746-786	786-∞
التكرار	13	36	38	13

أختبر فرض العدم أن زمن الحياة X للخلية الضوئية يتبع توزيعاً طبيعياً .

-١٤- قام مسئول مراقبة الجودة في مصنع لإنتاج وحدات معينة باختبار 200 وحدة من ناحية عدد السطوح التالفة وتم الحصول على البيانات التالية :

عدد السطوح التالفة	0	1	2	3	4	5	6فأكثر
التكرار المشاهد	90	62	31	13	3	1	0

استخدم توزيع بواسون لتوفيق هذه البيانات واختبر جودة التوفيق عند مستوى معنوية $\alpha = 0.05$.

-١٥- أوضحت الدراسات السابقة في مصنع للملابس أن 14% من المبيعات كانت للأشخاص الذين عمرهم أقل من 16 سنة و 38% كانت للأعمار من 16 إلى 20 سنة و 26% كانت للأعمار من 21 إلي 25 سنة و 22% كانت للأشخاص أكبر من 25 سنة. الجدول التالي يعطي النتائج التي تم الحصول عليها من عينة من 200 شخص .

العمـــر	اقل من 16	16-20	21-25	أكبر من 25
التكرار	22	62	60	56

ما هو فرض العدم الذي تتوقع أن تختبره ؟ ضع استنتاجك عند مستوى معنوية $\alpha = 0.05$.

-١٦- إذا كان X يمثل الزمن اللازم (بالأيام) لتصليح مكون ما في طائرة . المطلوب اختبار ما إذا كان توزيع بواسون بمتوسط $\mu = 3$ أيام نموذج مناسب لهذا المتغير وذلك باستخدام البيانات في الجدول التالي :

زمن التصليح (بالأيام)	0	1	2	3	4	5	6	≥ 7
التكرار	1	3	7	6	10	7	6	0

-١٧- يعطى الجدول التالي سرعة الإشعاع radial relocities لعدد 80 من النجوم الساطعة والمطلوب اختبار ما إذا كان التوزيع الطبيعي نموذج قياسي لهذا المتغير .

حدود الفئة	(-80,-70)	(-70,-60)	(-60,-50)	(-50,-40)	(-40,-30)
التكرار	1	2	2	2	8
حدود الفئة	(-30,-20)	(-20,-10)	(-10, 0)	(0,10)	(10,20)
التكرار	24	26	11	2	1
حدود الفئة	(20,30)				
التكرار	1				

-١٨- في عينة عشوائية من 1000 شخص وجد أن 452 يعطون صوتهم لمرشح ما. المطلــــوب اختبار فرض العدم أن 50% من الأصوات تذهب إلى المرشح وذلك عنـــــد مســـتوى معنويـــة $\alpha = 0.01$.

-١٩- خلال 1000 ساعة طيران لخمسين طائرة تم تسجيل عدد الطائرات m_x التي حدث لهـــا فشل في x مكون . النتائج في الجدول التالي :

x	0	1	2	3	4	≥ 5
m_x	7	10	9	7	6	11

اختبر فرض العدم أن المتغـــير X يتبـــع توزيـــع بواســـون بمعلمـــة $\mu = 2$ عنـــد مســـتوى معنوية $\alpha = 0.01$.

-٢٠-إذا كان عدد الصناديق المرفوضة من الحجم الواحد من مسحوق منظف من الأنـــواع A, B, C, D, E, F هي :

النوع	A	B	C	D	E	F
عدد الوحدات المرفوضة	270	308	290	312	300	320

أختبر فرض العدم أن الاختلاف بين الأنواع المختلفة يرجع إلى الصدفة باستخدام مستوى معنوية $\alpha = 0.05$.

-٢١- في بحث ميداني عن التفضيل لخمسة أنواع من مســـــاحيق التنظيـــف A,B,C,D,E تم الحصول على البيانات التالية :

النوع	A	B	C	D	E
التكرار	187	221	193	204	195

هل تعتقد أن المستهلكين غير قادرين على التمييز بين المساحيق الخمسة ؟ عند مستوى معنويـــــة $\alpha = 0.05$.

-٢٢- في مشكلة جينية يعتقد أن اللون البني يحدث باحتمال 25. والأبيـــض باحتمـــال 0.25 والمنقط باحتمال 0.5 .

(أ) اختبر فرض العدم أن هذا النموذج صحيح عند مســـــتوى معنويـــة $\alpha = 0.01$ إذا تم الحصول على البيانات التالية :

اللون	البني	الأبيض	المنقط
التكرار	5	15	20

(ب)أختبر فرض العدم أن الاحتمالات هي $\frac{1}{9}$ للبنى و $\frac{4}{9}$ للأبيض و $\frac{4}{9}$ للمنقط وذلك عند مستوى معنوية $\alpha = 0.05$.

-٢٣- إذا كان احتمال أن لاعب كرة السلة يصوب نحو الهدف هو 0.3 فإذا لعب 100 محاولة فحصل على 20 هدف أختبر فرض العدم $H_0: P = 3$ ضد الفرض البديل $H_1: P \neq 0.3$ حيث P هو احتمال التصويب نحو الهدف عند مستوى معنوية $\alpha = 0.01$.

-٢٤- اختيرت عينه عشوائية من 30 فرداً في جامعة ما وتم تصنيفهم تبعاً للجنس وعدد ساعات مشاهدة التليفزيون في خلال أسبوع. البيانات التي تم الحصول عليها معطاة في الجدول التالي :

المشاهدة	الجنس	
	ذكر	أنثي
أكثر من 25 ساعة	5	9
أقل من 25 ساعة	9	7

هل يمكن القول أن هناك علاقة بين الجنس وعدد ساعات مشاهدة التليفزيون وذلك عند مستوى معنوية $\alpha = 0.05$.

-٢٥- أخذت عينة عشوائية من 200 رجل متزوج وتم تصنيفهم في الجدول التالي تبعاً للتعليم وعدد الأطفال :

التعليم	عدد الأطفال		
	0-1	2-3	أكثر من 3
بسيط	14	37	32
متوسط	19	42	17
جامعي	12	17	10

أختبر فرض العدم أن عدد الأفراد في الأسرة غير مرتبط بالتعليم وذلك عند مستوى معنوية $\alpha = 0.01$.

-٢٦- لدراسة العلاقة بين امتلاك سيارة وامتلاك تليفزيون في بلد ما تم إجراء استطلاع للرأي على عينة من 20,000 شخص وقد تم الحصول على البيانات التالية :

عدد التليفزيونات التي يمتلكها الشخص	عدد السيارات التي يمتلكها الشخص			
	لا يوجد	1	2	3

←

لا يوجد	1000	900	100	2000
واحد	1500	2600	500	4600
أثنين أو اكثر	500	2500	400	3400

أختبر فرض العدم أن عدد السيارات التي يمتلكها الشخص مستقلة عن عدد التليفزيونات التي يمتلكها وذلك مستوى معنوية $\alpha = 0.05$.

-٢٧- أجري استطلاع للرأي على 300 طالب في ثلاث كليات للتجارة (في ثلاث جامعات مختلفة) وذلك لمعرفة المجال التجاري الذين يرغبون العمل فيه بعد التخرج وكانت الإجابات كالتالي :

المجال	A	B	C
الاستثمار	20	2υ	10
البنوك	15	25	30
الشركات	60	40	50
الاستشارات	5	15	10

يرغب المسئول عن البحث في معرفة ما إذا كان المجال الذي يعمل فيه الشخص مستقل عن الجامعة التي يتخرج منها وذلك عند مستوى معنوية $\alpha = 0.05$.

-٢٨- في عينة عشوائية من 174 شخص يقودون سيارة جديدة في مدينة ما تم تصنيفهم حسب عدد السيارات التي اشتراها كل قائد سيارة خلال العشرة سنوات السابقة والأعمار المختلفة لهم وذلك في الجدول التالي :

عدد السيارات	العمـــر		
المشتراة	اقل من 30	30 - 40	40 فأكثر
0	49	5	12
1	23	13	12
2	11	18	7
أكثر من 2	7	12	5

قدر ما إذا كان هناك علاقة بين العمر وعدد السيارات المشتراة .

-٢٩- الجدول التالي يبين 94 شخصاً مقسمين حسب التدخين والتعليم .

التدخين	متعلم	غير متعلم

←

يدخن	29	22
لا يدخن	24	19

المطلوب اختبار العلاقة بين التدخين والتعليم عند مستوى معنوية $\alpha = 0.05$.

-٣٠- تقوم إحدى الشركات بإنتاج ثلاثة مستويات من منتج معين . فـإذا اختيرت عينة عشوائية مكونه من 540 وحدة من المصنع وتم تسجيل عدد الوحدات التالفة في كل مستوى في الجدول التالي :

مستوى الإنتاج	نـوع الإنتاج	
	سليم	معيب
A	185	20
B	199	24
C	98	14

اختبر فرض العدم بعدم وجود علاقة بين مستوى الإنتاج ونوع المنتج (سليم ومعيب) وذلك عند مستوى معنوية $\alpha = 0.05$.

-٣١- أجرى بحث ميداني لتقدير ما إذا كان هناك علاقة بين الاتجاه السياسي وغزو الفضاء وتم الحصول على البيانات التالية :

الاتجاه السياسي	التدعيـم		
	جيد جداً	جيـد	لا يدعم
جمهوري	8	12	10
ديمقراطي	10	17	10
مستقل	12	6	12

أختبر فرض العدم بعدم وجود علاقة بين الاتجاه السياسي وغزو الفضاء .

-٣٢- أراد مسئول في مكتبة لبيع الكتب العلمية دراسة ما إذا كان هناك علاقة بين لون الإعلانات المرسلة إلى العملاء وكمية المبيعات . أرسل 200 إعلان إلى مجموعة من العملاء وتم الحصول على البيانات التالية :

	إعلان ملون	إعلان غير ملون
الكمية المباعة	70	20
الكمية الغير مباعة	30	80

المطلوب دراسة العلاقة بين لون الإعلان وكمية المبيعات عند مستوى معنوية $\alpha = 0.05$.

-٣٣- يعطى الجدول التالي التقديرات التي حصل عليها 272 طالباً في اختبارين مختلفين.

الاختبار الأول	الاختبار الثاني		
	ممتاز	جيد	مقبول
ممتاز	9	9	99
جيد	29	7	39
مقبول	59	2	19

هل يمكن الجزم بعدم وجود ارتباط بين درجات الاختبارين عند مستوى معنوية $\alpha = 0.01$.

-٣٤- في عينة عشوائية من 750 شخص تم تصنيفهم حسب الدخل والوزن في الجدول التالي :

الوزن	الدخل		
	منخفض	متوسط	عالي
نحيف	100	50	50
متوسط	50	200	70
بدين	120	60	50

أختبر فرض العدم أن العاملين (الدخل والوزن) مستقلين عند مستوى معنوية $\alpha = 0.05$.

-٣٥- الجدول التالي يعطى عدد الوحدات التالفة السليمة المنتجة بواسطة آلتين .

	وحدات تالفة	وحدات سليمة
الماكينة A	25	375
الماكينة B	42	558

هل عدد الوحدات التالفة مستقل عن نوع الماكينة التي تنتجها. وذلك عند مستوى معنوية $\alpha = 0.05$.

-٣٦- أعطيت لعينة من 100 شخص مسحوق للغسيل قياسي وذلك لتسجيل ما إذا كان المسحوق فعال أو غير فعال ثم أعطى لهم مرة أخري مسحوق جديد وتم سؤالهم مرة أخرى عن رأيهم في المسحوق . نتائج الاستقصاء في الجدول التالي :

المسحوق القياسي	المسحوق الجديد	
	فعال	غير فعال
فعال	20	10
غير فعال	50	20

هل يمكن القول بأن هناك ارتباط بين الفاعلية ونوع المسحوق ؟ وذلك عند مستوى معنوية $\alpha = 0.01$.

-٣٧- يعطى الجدول التالي تصنيف لعينة عشوائية من 2764 شخص حسب الدخل بالدولار والفترة منذ أخر زيادة لاستشارة طبيب .

الدخل	منذ 6 شهور	من 7 شهور لسنة	أكثر من سنة
أقل من 3000	186	38	35
3000-4999	227	54	45
5000-6999	219	78	78
7000-9999	355	112	140
أكثر من 10,000	653	285	259

المطلوب اختبار ما إذا كان هناك استقلال بين المتغيرين (الدخل وزيارة الطبيب) وذلك عند مستوى معنوية $\alpha = 0.05$.

-٣٨- قامت إحدى الشركات بعمل دورات تدريبيه للعاملين بها لرفع كفاءتهم . اختيرت عينة عشوائية من 100 عامل وتم تصنيفهم تبعاً لكفاءتهم قبل وبعد عمل الدورات التدريبية والنتائج معطاة في الجدول التالي :

قبل الدورات	بعد الدورات		المجموع
	غير ماهر	ماهر	
غير ماهر	20	80	100
ماهر	2	98	100
المجموع	22	178	200

هل يمكن القول أن هناك علاقة بين التدريب والكفاءة ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٣٩- قام صاحب متجر بتسجيل المعلومات التالية عن 450 شخص من المترددين على المتجر والبيانات معطاة في الجدول التالي :

	ذكور	إناث
يشتري	60	60
لا يشتري	150	180

هل يمكن القول أن الشراء يرتبط بالجنس ؟ وذلك عند مستوى معنوية $\alpha = 0.01$.

٤٠- لدراسة العلاقة بين حساسية الجلد من ضوء الشمس ولون العين حصل طبيب متخصص في الأمراض الجلدية على البيانات التالية وذلك من عينة عشوائية من 100 شخص .

لون العين	تأثير الأشعة		
	قوي	متوسط	ضعيف
أزرق	19	27	4
رمادي أو اخضر	7	8	5
بني	1	13	16

هل يمكن القول أن هناك علاقة بين لون العين وحساسية الجلد عند مستوى معنوية $\alpha = 0.05$.

٤١- أعطي دواء لثلاثة مجموعات وسجلت النتائج في الجدول التالي :

الدواء	نسبة المرضى الذين تم شفائهم	نسبة المرضي الذين لم يشفوا
A	95	35
B	90	10
C	85	15

قدر ما إذا كانت هذه البيانات توضح اختلاف استجابة الشفاء للأدوية الثلاثة وذلك عند مستوى معنوية $\alpha = 0.05$

٤٢- يقوم مصنعين بإنتاج نفس المنتجات A,B,C الجدول التالي يعطى الوحدات المنتجة من كل مصنع .

المصنع	المنتج			المجموع
	A	B	C	
X	42	13	33	88
Y	20	21	58	99

هل يمكن القول بأن هناك فرق معنوية بين الإنتاج للمصنعين ؟ عند مستوى معنوية $\alpha = 0.05$

٤٣- في معهد للأورام تم تجربة علاج جديد يسمى Q 134 على عينة عشوائية من 170 مريض . أيضا اختيرت عينة عشوائية من 170 مريض وتم علاجهم بالعلاج القديم المسمى Q133 والنتائج معطاة في الجدول التالي :

العلاج	شفاء	وفاة
Q134	150	20
Q133	130	40

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_j ; j = 1,2$ وذلك عند مستوى معنوية $\alpha = 0.05$.

٤٤- اختيرت عينات عشوائية من الطلبة في فرق دراسية مختلفة وتم حساب درجاتهم وتسجيل تقديراتهم في الجدول التالي :

الفرقة	التقدير		
	ضعيف	جيد ومقبول	جيد جدا وممتاز
الفرقة الأولى	24	58	18
الفرقة الثانية	36	112	52
الفرقة الثالثة	80	230	90
الفرقة الرابعة	60	200	40

هل تدل هذه البيانات على أن توزيع الطلاب حسب تقديراتهم يعتمد على الفرقة الدراسية الموجود فيها الطالب ؟ استخدم مستوى معنوية $\alpha = 0.05$.

٤٥- يوجد في شركة سنترال A وسنترال B ويعتقد العامل على السنترال A أنه يستقبل مكالمات دولية أكثر من السنترال B لاختبار هذا الفرض تم الحصول على البيانات التي في الجدول التالي :

السنترال	المكالمات المحلية	المكالمات الدولية
A	400	100
B	328	72

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_j ; j = 1,2$ وذلك عند مستوى معنوية $\alpha = 0.05$.

٤٦- قام صاحب سيارات بعمل استطلاع للرأي عن تفضيل شراء نوع جديد من السيارات وذلك في ثلاثة مدن A,B,C . اختيرت عينات عشوائية من الثلاثة مدن وتم سؤالهم " هل تفضلون شراء السيارة الجديدة " ؟ وتم الحصول على البيانات التالية :

نتائج الاستطلاع	المدن		
	A	B	C
يفضل الشراء	100	160	190
لا يفضل الشراء	50	40	60

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_j ; j = 1,2$.

-٤٧- في استطلاع للرأي على الأشخاص بدون عمل في أربعة مناطق في بلد ما اختيرت أربع عينات من الحجم 300 واحدة من كل منطقة وتم تصنيفهم حسب المدة التي مكثوا فيها بدون عمل :

المدة بدون عمل بالأسبوع	المناطق				المجموع
	NE	SE	Nw	Sw	
أكثر من 20	80	75	60	65	280
من 10-20	115	105	110	90	420
أقل من 10	105	120	130	145	500

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_{j|4} = P_j ; j = 1,2,3$

-٤٨- بفرض أن جدول توافق من نوع 2 x 3 لمقارنة عينتين مستقلين . بفرض أن عينة عشوائية من الرجال وعينة عشوائية أخرى من النساء أعطوا رأيهم في بحث ميداني وتم تسجيل النتائج في الجدول التالي :

	يوافق	لا يوافق	لم يقرر
سيدة	118	62	25
رجل	84	78	37

هل يمكن القول أن الرجال والسيدات يفكرون بطريقة مختلفة ؟

-٤٩- إذا تم تصنيف أنواع الفشل في طائرة إلى كهربائي - ميكانيكا - غير ذلك . فإذا تم تصميم نوعين من الطائرات I,II. المطلوب اختبار فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_j ; j = 1,2,3.$ وذلك عند مستوى معنوية $\alpha = 0.05$

وذلك بالاعتماد على البيانات التالية :

	ميكانيكي	كهربائي	غير ذلك
التصميم I	50	30	60
التصميم II	40	30	40

-٥٠- أجريت دراسة على عادة التدخين بين الذكور والإناث في منطقة ما. اختيرت عينتين واحدة من 200 رجل والأخرى من 200 سيدة والبيانات في الجدول التالي :

	يدخن	لا يدخن
رجال	110	90

←

سيدات	104	96

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_j \, , j = 1,2.$ وذلك عند مستوى معنوية $\alpha = 0.05$

-٥١- يعتقد شخص أن نسبة وجود صفة ما في ثلاث مجتمعات واحدة وتساوى 0.3. اختيرت عينات عشوائية من الحجم 50, 100, 50 من المجتمعات 1,2,3 على التوالي . يعطى الجدول التالي البيانات التي تم الحصول عليها في كل عينة . أختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_j \, , j = 1,2.$

	نعم	لا	المجموع
العينة 1	20	30	50
العينة 2	25	75	100
العينة 3	25	25	50

-٥٢- بفرض أن لديك البيانات التالية في بحث ميداني :

	نعم	لا	المجموع
العينة 1	10	40	50
العينة 2	50	20	100
العينة 3	30	20	50

إذا كان يعتقد أن نسبة الذين يقولون نعم هو P_1=0.8 اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_j , j = 1,2$

-٥٣- في بحث ميداني تم تصنيف ثلاثة عينات من المزارعين I, II, III حسب نوع العمل في الأرض التي يزرعها (خليط – بأجر – يمتلك) . البيانات معطاة في الجدول التالي :

	يمتلك	بأجر	خليط
العينة I	36	67	49
العينة II	31	60	49
العينة III	58	87	80

أختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_{1|3} = P_j , j = 1,2,3$ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٥٤- يتكون نظام من أربعة أجزاء تعمل مستقلة عن بعضها فإذا كان P_i يرمز إلى احتمال أن يعمل الجزء i بنجاح . أختبر فرض العدم $P_1=.9$, $P_2=.9$, $P_3=0.8$, $P_4=.8$ وذلك باستخدام البيانات في الجدول التالي والتي تعطي عدد مرات النجاح لكل جزء في 50 محاولة.

الجزء	1	2	3	4
عدد مرات النجاح	40	48	45	40

-٥٥- البيانات التالية تمثل 20 مفردة تم توليدها على الحاسب الآلي :

.21	.27	.92	.91	.86	.31	.29	.1	.48	.81
.70	5.9	.51	.97	.81	.84	.47	.39	.39	.31

(أ) أختبر فرض العدم $H_0 : M = .5$ ضد الفرض البديل $H_1: M > .5$ وذلك عند مستوى معنوية $\alpha = 0.05$.

(ب) أختبر فرض العدم $H_0 : M = .21$ ضد الفرض البديل $H_1: M > .21$ وذلك عند مستوى معنوية $\alpha = 0.01$.

-٥٦- البيانات التالية تمثل الدخل السنوي لعشرين أسرة بالدولار

5850	13670	14350	48160	2370	19340	30700	20410	25570	20130
45150	13570	12110	21660	27620	17880	24840	496	19340	30700

أختبر فرض العدم $H_0 : M = 24800$ ضد الفرض البديل $H_1: M < 24800$ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٥٧- إذا كانت أوزان إناث القرود بالكيلو جرام لعينة من الحجم n=15 هي :

10.05	9.80	9.25	9.10	8.40	8.75	9.60	9.50	8.3
			9.30	9.20	9.80	9.60	10.00	8.15

أختبر فرض العدم $M = 8.81$ ضد الفرض البديل $H_1 : M > 8.41$ وذلك باستخدام اختبار الإشارة .

-٥٨- تم سؤال 14 شخص من المدخنين عن العمر الذي بدءوا التدخين عنده وكان لدينا النتائج الآتية :

الشخص	1	2	3	4	5	6	7
العمر	22	25	37	28	15	14	22
الشخص	8	9	10	11	12	13	14
العمر	16	18	17	23	16	20	18

هل يمكن اختبار الفرض القائل بأن وسيط المجتمع المسحوبة منه هذه العينة يساوى 20 ضد الفرض القائل أن الوسيط لا يساوى 20 وذلك عند مستوى معنوية $\alpha = 0.05$.

-٥٩- إذا كانت M ترمز لوسيط توزيع متماثل من النوع المتصل استخدم اختبار الإشارة لاختبار فرض العدم $H_0: M = 75$ ضد الفرض البديل $H_1: M > 75$ وذلك عند مستوى معنوية $\alpha = 0.05$ وذلك بالاعتماد على عينة عشوائية من الحجم $n=18$. المشاهدات في الجدول التالي تعطى انحرافات مشاهدات العينة عن 75 أي (x_i-M) , $i=1,2,\dots,18$

1.5	-0.5	1.6	0.4	2.3	-0.8	3.2	0.9	2.9
0.3	1.8	1.5	-0.1	1.2	2.5	0.6	-0.7	1.9

-٦٠- البيانات التالية تمثل أطوال عينة من 16 شجرة بالسنتيمترات :

122	117	127	102	137	135	124	145
130	139	127	129	143	113	116	135

أختبر ما إذا كان وسيط المجتمع هو $M = 130$ وذلك عند مستوى معنوية $\alpha = 0.05$.

(الفرض البديل $M \neq 130$) .

-٦١- أستخدم اختبار الإشارة لتحليل فئة الأزواج في الجدول التالي وذلك لاختبار فرض العدم $H_0 : M_D = 0$ ضد الفرض البديل $H_1: M_D > 0$ وذلك عند مستوى معنوية $\alpha = 0.05$.

الشخص	1	2	3	4	5	6	7	8	9	10
المعالجة A	46	41	37	32	28	43	42	51	28	27
المعالجة B	32	43	37	32	31	39	44	53	26	31

-٦٢- البيانات في الجدول التالي تمثل المبيعات لمشروب بارد A في يوم ما التي اشتراها 12 شخص وأيضا المبيعات لنفس الأشخاص من المشروب B في نفس اليوم .

المشاهدة	المشروب A	المشروب B
1	4	5
2	6	5
3	3	3
4	5	6
5	7	8
6	8	6
7	9	6
8	3	10
9	4	4
10	4	5
11	5	5

12	4	6

أختبر فرض العدم $H_0 : M_D = 0$ ضد الفرض البديل $H_1: M_D > 0$ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٦٣- أخذت عينة من 10 أطفال بإحدى المدارس ودونت أوزانهم ثم أعطي كل منهم وجبة غذائية وذلك لمدة 4 شهور متتالية ثم دونت أوزانهم فكانت النتائج كالآتي :

الطفل	1	2	3	4	5	6	7	8	9	10
قبل الوجبة	130	127	129	141	137	134	139	140	138	139
بعد الوجبة	129	124	139	177	134	136	137	135	133	132

المطلوب اختبار أثر تعاطي الوجبة في زيادة وزن مجتمع هؤلاء الأطفال عند مستوى معنوية $\alpha = 0.05$.

-٦٤- استخدمت طريقتين لتقدير مستوى مركب Protein Bound iodine وذلك باستخدام 10 أنثى بالغة. استخدم اختبار الإشارة في تقدير ما إذا كانت النتائج في الجدول التالي والمقدرة بالطريقة B أكثر معنوية من القيم المقدرة بالطريقة A .

الشخص	1	2	3	4	5	6	7	8	9	10
A	4	6	3	5	7	8	9	3	4	5
B	5	5	3	6	8	6	10	4	5	5

-٦٥- اختير 18 زوج (توائم) من ذكور حيوان ما حيث أعطى الغذاء I لواحد من التوائم وأعطى الغذاء II للآخر في كل زوج . الأوزان خلال 8 أسابيع معطاة في الجدول التالي :

I	111	102	90	110	108	125	125	99	121
II	101	97	90	96	95	110	107	85	104
I	133	101	98	109	107	124	115	90	120
II	119	98	97	104	90	109	106	84	105

استخدم اختبار الإشارة لاختبار الفرض بعدم وجود فرق معنوي بين الغذائيين ضد الفرض البديل أن الغذاء I أحسن وذلك عند مستوى معنوية $\alpha = 0.05$.

-٦٦- البيانات التالية تعطى عدد ضربات القلب في الدقيقة لعشرة فئران مرة عند وضع الفأر في قفص بدون رفاق ومرة أخرى في وجود رفاق له .

الفأر	1	2	3	4	5	6	7	8	9	10

463	462	462	456	450	426	418	415	409	402	وحيد
523	494	461	461	535	476	454	484	470	437	الفأر مع رفاق

أختبر فرض العدم $H_0: M_D = 0$ ضد الفرض البديل $H_1: M_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

٦٧- أجريت مسابقة لعمل كيك ذات مواصفات معينة لمجموعتين من السيدات أ و ب وكانت الدرجات التي حصلت عليها كل سيدة معطاة في الجدول التالي :

91	92	96	97	97	93	92	90	المجموعة أ
91	90	91	87	94	95	88	89	المجموعة ب

هل يمكن القول أن العينتين تم اختيارهما من نفس المجتمع ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

٦٨- يعطى الجدول التالي عدد الصمامات الكهربية المنتجة في خطين من خطوط الإنتاج A,B وذلك خلال فترة 10 أيام .

A	170	164	140	184	174	142	191	169	161	200
B	201	179	159	195	177	170	185	179	170	212

هل يمكن القول أن العينتين تم اختيارهما من نفس المجتمع وذلك عند مستوى معنوية $\alpha = 0.05$.

٦٩- نوعين من البلاستيك ، كل نوع منتج بطريقة مختلفة . وقد تم اختبار ultimate strength لكل مفردة في عينة عشوائية من النوع الأول وأيضا عينة عشوائية من النوع الثاني (وحدة القياس 1,000 رطل لكل بوصة مربعة).

15.3	18.7	22.3	17.6	19.1	14.8	النوع 1
21.2	22.4	18.3	19.3	17.1	27.7	النوع 2

هل يمكن القول أن النوعين لهما نفس ultimate strength وذلك عند مستوى معنوية $\alpha = 0.05$.

٧٠- لمقارنة معدل النبض في ذكور الفئران بمعدل النبض لإناث الفئران تم الحصول على البيانات التالية :

الذكور	74	77	78	75	72	71
الإناث	60	83	73	84	82	79

هل يمكن القول أن معدل النبض واحد في الذكور والإناث وذلك عند مستوى معنوية $\alpha = 0.05$.

-٧١- يعطي الجدول التالي أزمنة الحياة لكرات التحميل وذلك باستخدام اختبارين مختلفين . المطلوب اختبار فرض العدم H_0 : المجتمعين لهما نفس التوزيع ضد الفرض البديل H_1 : قيم x'_s تتجه لأن تكون أقل من قيمة y'_s وذلك عند مستوى معنوية $\alpha = 0.05$.

الاختبار الأول	140.3	158.0	183.9	132.9	117.8	98.7	164.8	93.4
الاختبار الثاني	193.0	172.5	173.3	204.7	172.0	152.9	216.0	422.6

-٧٢- لمقارنة القلوية في بركتين (مقاسه بالمليجرام لكل لتر) ، أخذت عينتين من الحجم 5 من كل بركة . استخدم اختبار Wilcoxon لتقدير ما إذا كان هناك فرق معنوي في القلوية بين البركتين :

البركة I	102	116	122	112	104
البركة II	108	117	115	120	105

-٧٣- إذا كانت كمية النيكوتين لنوعين من السجائر مقاس بالمليجرام كالآتي :

النوع A	22.1	24.0	26.3	24.8	25.4	26.1	23.3	22.1	24.1	22.3
النوع B	24.1	20.6	23.1	22.5	24.0	26.2	21.6	22.2	21.9	25.4

هل يمكن القول أن كمية النيكوتين واحدة في النوعين وذلك عند مستوى معنوية $\alpha = 0.05$.

-٧٤- إذا كان عدد البكتريا في وحدة حجم وذلك لأربع وحدات من الحجم في المزرعة A وأيضا في المزرعة B البيانات معطاة في الجدول التالي . اختبر فرض العدم H_0 : أن المجتمعين لهما نفس التوزيع ضد الفرض البديل H_1: قيم x'_s تتجه لأن تكون أكبر من قيم y'_s :

A	27	361	26	25
B	32	29	35	29

-٧٥- استخدم اختبار Wilcoxon لتقدير ما إذا كان هناك فرق معنوي بين الفئتين التاليين من البيانات :

A	82	86	30	21	38	29	29	19
B	124	116	54	54	110	29	39	54

-٧٦- في مصنع للورق يوجد نوعين من الورق A,B . أخذت عينة عشوائية من 5 لفات من كل نوع وقيست قوة التمزق لمفردات كل عينة . البيانات معطاة في الجدول التالي :

النوع A	154	143	135	140	128
النوع B	149	162	160	154	175

هل يمكن القول أن قوة التمزق مختلفة في النوعين وذلك عند مستوى معنوية $\alpha = 0.05$.

-٧٧- لمعرفة تأثير طريقة حديثة لزراعة الذرة استخدمت إحدى محطات التجارب الزراعية هذه الطريقة الحديثة في زراعة 8 قطع وقورن المحصول الناتج مع المحصول الناتج من 9 قطع متجاورة زرعت بالطريقة العادية وكانت النتائج (عدد الشجيرات بالفدان كالآتي):

الطريقة العادية	40.1	28.7	33	36.9	34.9	33.5	37.1	40.2	
الطريقة الحديثة	55.1	49.2	23.4	46.5	52	52.2	51.3	40.9	50.1

هل هذه النتائج تعطى دليلاً على أن الطريقة الحديثة أحسن من الطريقة العادية .

-٧٨- تعطى البيانات في الجدول التالي المبيعات اليومية لمحلية تجاريين ، العينة الأولى للمحل 1 تعطى المبيعات اليومية خلال 12 يوماً والعينة الثانية تعطي المبيعات اليومية للمحل 2 خلال 10 أيام . أختبر فرض العدم H_0: المجتمعين لهما نفس التوزيع ضد الفرض البديل H_1: قيم x'_s تتجه لأن تكون أكبر من قيم y'_s وذلك عند مستوى معنوية $\alpha = 0.01$. المبيعات مقاسة بالدولار.

المحل 1	100	125	60	137	82	99	150	98	143	122	95	110
المحل 2	82	98	115	143	65	123	128	93	135	120		

-٧٩- اختبرت مجموعتان من الأرانب الأولى من 10 أرنبا أعطيت الغذاء A والثانية من 12 أرنباً أعطيت غذاء B وكانت الزيادة في الوزن بعد فترة معينة هي :

A	24	34	30	30	33	22	12	23	27	29		
B	30	36	38	33	38	34	30	35	30	36	38	36

هل يمكن القول أن العينتين تم اختيارهما من نفس المجتمع ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨٠- يعطى الجدول التالي درجات الذكاء لمجموعتين من الطلبة والمطلوب اختبار هل العينتين تم اختيارهما من نفس المجتمع وذلك عند مستوى معنوية $\alpha = 0.05$.

I	111.7	119.5	118.7	111.2	117.2	117.4
II	116.6	115.8	115.8	115.4	115.1	115.0
I	117.4	116.8	115.0	114.1	114.1	112.2
II	113.9	112.4	112.7	111.9	112.8	110.2

أختبر فرض العدم أن المجتمعين لهما نفس التوزيع ضد الفرض البديل أن قيم x'_s تتجه لأن تكون أكبر من قيم y'_s وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨١- البيانات التالية تعطى كمية الوقود المستهلكة لثلاث أنواع من المحركات (الوحدات المستخدمة في القياس kilometers per liter).

النوع A	31	30	29	22	32	25			
النوع B	12	13	26	24	11	26	27		
النوع C	17	20	23	9	15	18	19	14	8	5

هل يوجد فروق معنوية بين الأنواع الثلاثة عند مستوى معنوية $\alpha = 0.05$.

-٨٢- يرغب مسئول في مصنع لإنتاج الورق في تقدير ما إذا كان هناك اختلاف معنوي في مقاومة الرطوبة لأربعة طرق لتخزين الورق . نتائج الاختبار لطرق التخزين الأربعة معطاة في الجدول التالي :

الطريقة			
A	B	C	D
5.3	4.3	6.0	5.6
5.3	3.7	5.0	8.0
6.5	3.8	5.6	5.4
5.4	4.6	4.9	6.5
7.6	4.1	4.5	8.5

هل يمكن القول أن هناك اختلاف في كمية الرطوبة بين طرق التخزين المختلفة ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨٣- قررت شركة تطوير ثلاثة أنواع من المحركات واختيار واحد منهم لسيارة فاخرة يراد تسويقها . لهذا الغرض أجريت تجربة لتقدير استهلاك الوقود لكل محرك حيث اختيرت عينة

عشوائية من كل نوع وتم اختبارهم في ظروف الطريق الطبيعي . البيانات التي تم الحصول عليها من الاختبار في الجدول التالي :

1	31	30	29	22	32	25				
2	12	13	26	24	11	16	27	10		
3	3	17	20	32	9	15	18	19	14	8

المطلوب اختبار فرض العدم H_0: توزيعات المجتمعات الثلاثة متطابقة ضد الفرض البديل H_1: التوزيعات ليس لها نفس الوسيط وذلك عند مستوى معنوية $\alpha = 0.01$.

-٨٤- لدراسة تأثير نوع جديد من سيرم الدم على وقف سرطان الدم ، أختير 9 فئران في مراحل متقدمة من المرض وتم تعويض 5 منهم للمعالجة و 4 لم يتم تعويضهم . يعطى الجدول التالي أزمنة الحياة ، بالسنة ، من بداية التجربة .

عولج	2.1	5.3	1.4	4.6	0.9
لم يعالج	1.9	0.5	2.8	3.1	

هل يمكن استنتاج أن سيرم الدم له تأثير فعال على إيقاف سرطان الدم ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨٥- البيانات التالية تعطي أعمار البطاريات (بالساعة) لثلاثة أنواع من البطاريات المستخدمة في الآلات الحاسبة المحمولة .

I	II	III
12	17	21
19	25	15
19	13	16
16	31	29
20	36	32
24	15	14
30	20	18
15	35	26

هل يوجد اختلاف في متوسط أعمار البطاريات للأنواع الثلاثة ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨٦- في تجربة تم تقدير مستويات مركب antecubital vein cortisol لثلاث مجموعات من المرضى . البيانات معطاة في الجدول التالي :

I	262	307	211	323	454	339	304	154	287	326
II	465	501	455	355	468	360				
III	343	772	207	1048	838	687				

المطلوب اختبار فرض العدم أن توزيعات المجتمعات الثلاثة متطابقة ضـــد الفـرض البديــل أن توزيعات المجتمعات الثلاثة ليس لها نفس الوسيط وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨٧- قام مسئول بتسجيل الإنتاج اليومي لثلاثة ماكينات وحصل على البيانـــات المعطـــاة في الجدول التالي :

الآلة الأولي	100	110	92	95	108
الآلة الثانية	99	97	90	101	98
الآلة الثالثة	98	104	113	97	103

أختبر فرض العدم أن الماكينات الثلاثة تعطي نفس الإنتاج في اليوم وذلك عند مستوى معنويـــة $\alpha = 0.05$.

-٨٨- زرعت ثلاثة أنواع من الأسمدة عشوائية على مجموعة من قطـــع الأراضـــي المتجـــاورة والمزروعة بنوع واحد من محصول الذرة فكان المحصول الناتج كما يلي :

السماد الأول 83, 80, 76, 77, 85, 62

السماد الثاني 69, 50, 74, 71, 73, 68

السماد الثالث 55, 77, 65, 61, 80, 82

أختبر عند مستوى معنوية $\alpha = 0.01$ أن تأثير السماد متساوي ضد الفرض البديل أنها ليســت جميعاً متساوية .

-٨٩- للمقارنة بين ثلاثة أنواع من الأدوية في إنقاص الوزن تم اختيار عينة عشـــوائية مـــن 21 شخصاً وقسمت إلي ثلاثة مجموعات حيث أعطى لكل مجموعة نوع من هذا الدواء وبعد ثلاثـــة أشهر سجل النقص في أوزانهم بالكيلو جرام كالآتي :

الدواء

A 4.1, 5.2, 5.3, 5.6, 7.1, 7.2, 8.0

B 2.3, 2.4, 3.2, 3.2, 3.4, 3.5, 3.6

C 6.2, 6.9, 7.0, 7.8, 8.8, 9.0, 9.5

أختبر فرض العدم أن تأثير الأنواع الثلاثة متماثل ضد الفرض البديل أن هذا التأثير يختلف مـــن نوع إلى آخر عند مستوى معنوية $\alpha = 0.05$.

-٩٠- البيانات التالية تمثل الإنتاج اليومي من الأكواب الزجاجية لثلاث ماكينات وذلك مـــن عينة عشوائية من 12 يوماً. أختبر معنوية الفروق بين الماكينات الثلاثة وذلك عند مستوى معنوية $\alpha = 0.05$.

A 340, 345, 330, 342, 338

B 339, 333, 344

C 347, 343, 349, 355

-٩١- فيما يلي مستوى مركب Protoporphyrin في الدم لمجموعتين من المرضى ومجموعة من الأصحاء (ملليجرام لكل 100سم) .

مجموعة الأصحـــاء	21, 25, 46, 29, 37, 77, 27, 57, 71, 55, 30, 25, 45, 36, 29
المجموعة الأولي من المرضى	77, 171, 285, 81, 452, 511 173, 914, 83, 152, 770
المجموعة الثانية من المرضى	37, 28, 39, 44, 45, 28, 37, 19, 67, 11, 36, 8, 75, 147, 10

فهل تدل هذه البيانات على اختلاف مستوى مركـــب Protoporphyrin في الـــدم لهـــذه المجموعات الثلاثة عند مستوى معنوية $\alpha = 0.05$.

-٩٢- البيانات التالي تمثل نسبة غدة الأدرينالين لكل جرام من وزن الجسم في ثلاثة أنواع مـــن الفئران :

المجموعة الأولي	85.4	127.79,	137.7,	139.5,	139,	155.1
المجموعة الثانية	180.3,	192.2,	190.6,	233.0,	220.5	
المجموعة الثالثة	138.6,	140.8,	144.7,	183.3,	206.6	

اختبر معنوية الفروق بين الأنواع الثلاثة عند مستوى معنوية $\alpha = 0.05$.

-٩٣- في دراسة لمقارنة مستوى الكوليسترول في ثلاثة مجاميع كانت النتائج كمايلى :

المجموعة الثانية	262, 307, 200, 320, 453, 339, 303 153, 288, 355
المجموعة الثانية	464, 500, 454, 354, 467, 361
المجموعة الثالثة	342 771 206 1047 837 687

أختبر معنوية الفروق بين المجموعات الثلاثة عند مستوى معنوية $\alpha = 0.05$.

-٩٤- البيانات التالية تمثل عدد الأسماك التي تم صيدها سنويا عند موقع خاص في بحيرة كبيرة ، البيانات مقاسه بالطن . حلل هذه البيانات لتقدير ما إذا كان التغير السنوي عشوائياً أم لا ؟

السنة	1943	1944	1945	1946	1947	1948	1949	1950	1951
العدد	56	57	61	59	61	51	55	52	48
السنة	1952	1953	1954	1955	1956	1957	1958	1959	1960
العدد	51	49	47	51	41	48	39	42	41

-٩٥- البيانات التالية تعطي عدد الحمام في منطقة معينة (محمية) خلال السنوات من 1954 حتى 1965 .

السنة	1954	1955	1956	1957	1958	1959
العدد	13	13	12	10	8	11
السنة	1960	1961	1962	1963	1964	1965
العدد	9	7	6	5	7	8

هل تعتقد أن حجم المجتمع يتأثر عشوائياً من سنة إلى أخري ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٩٦- البيانات التالية تمثل الأزمنة بين الحوادث في مدينة كبيرة

8.66 11.28, 10.43, 10.89, 11.49, 15.92, 12.5, 13.86, 13.32

أختبر فرض العدم أن العينة عشوائية ضد الفرض البديل أن العينة غير عشوائية عند مستوى معنوية $\alpha = 0.05$.

-٩٧- بفرض أن لدينا عينة من 25 شخص وكنا نرمز للشخص الذكر في العينة بالرمز M وبالرمز F للأنثى وكانت لدينا النتائج الآتية

MF MMMM F MMM FFFF M FF MMM FFFFF .

المطلوب عند مستوى معنوية $\alpha = 0.05$ تحديد هل العينة عشوائية أم لا ؟

-٩٨- يعطى الجدول التالي درجات حرارة موجبة وسالبة

9, 10, 11.5, 3, -1, 3, 4, -7, -6, 3,
-9, 5, 9, 6, 5, 2, 7, -2, -5, -4, -6

والمطلوب بمستوى معنوية $\alpha = 0.05$ معرفة هل هذه البيانات عشوائية بالنسبة للإشارة أم لا ؟

-٩٩- أجرى اختبار لمجموعة من الموظفين قبل تدريبهم وبعد 6 أشهر من التدريب وقد تم ترتيبهم تبعاً لجودة إنتاجهم . البيانات التالية تعطي الرتب التي حصلوا عليها .

الموظف	A	B	C	D	E	F	G	H	I	J	K	L
الرتب قبل التدريب	2	3	1	6	7	4	8	5	12	11	9	10
الرتب بعد التدريب	1	2	3	4	5	6	7	8	9	10	11	12

أوجد معامل سبيرمان .

-١٠٠- في وكالة لبيع السيارات أجريت دراسة على 15 موظف في قسم المبيعات وذلك لدراسة العلاقة بين درجة الاختبار التي حصل عليها الموظف عند تعينه وعدد السيارات المباعة خلال السنة الأولي من التعين

الشخص	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
x الدرجة	72	88.5	70	87	71	85	89	93	98	96	86	82	88	83	80
y عدد السيارات	314	422	322	440	287	415	463	497	510	512	432	390	453	374	385

أختبر فرض العدم H_0: المتغيرين مستقلين ضد الفرض البديل H_1: توجد علاقة بين المتغيرين في نفس الاتجاه وذلك عند مستوى معنوية $\alpha = 0.05$.

-١٠١- أجريت مقابلة شخصية من قبل شخصين A,B وذلك لثمانية أشخاص متقدمين لوظيفة ما وتم تسجيل الدرجات التالية :

الشخص	1	2	3	4	5	6	7	8
A	12	17	14	7	10	12	6	3
B	16	15	18	5	3	10	7	4

حول الدرجات إلى رتب وأحسب معامل سبيرمان .

-١٠٢- البيانات التالية تعطي الرتب التي حصل عليها 10 سكرتيرات يكتبن على الآلة الكاتبة أولاً في الظروف العادية ثم في ظروف اختبار . اوجد معامل سبيرمان

السكرتيرة	A	B	C	D	E	F	G	H	I	J
في ظروف الاختبار	1	2	3	4	5	6	7	8	9	10
في الظروف العادية	3	4	6	1	2	8	9	10	7	5

-١٠٣- قامت شركة بإجراء اختبار لـ 7 أشخاص من المتقدمين لوظيفة ما . الاختبار الأول مقابلة شخصية والآخر تحريري وتم الحصول على البيانات التالية :

الشخص	1	2	3	4	5	6	7
الرتبة للمقابلة الشخصية	4	1	7	6	2	3	5
الرتبة للامتحان التحريري	5	2	7	4	1	3	6

أوجد معامل سبيرمان .

-١٠٤- طلب أستاذ في الجامعة من 10 طلاب تقديم بحثين في الامتحان النهائي . الدرجـــات التي حصل عليها كل طالب هي :

الطالب	1	2	3	4	5	6	7	8	9	10
درجة البحث الأول (الدرجة من 100)	58	56	54	65	58	60	59	51	53	56
درجة البحث الثاني (الدرجة من 10)	6	7	5	8	10	5	7	6	2	4

أوجد معامل سبيرمان .

-١٠٥- الجدول التالي يسجل كمية المطر (اليومية) وعدد ساعات ظهور السحب (بالساعات) في بلد ما خلال 11 فترة زمنية والمطلوب إيجاد معامل سبيرمان للرتب بين سقوط الأمطار وظهور السحب.

الفترة	1	2	3	4	5	6	7	8	9	10	11
سقوط المطر	15.1	15.8	14.9	16.6	12.6	17.4	16.1	13.7	15.5	17.5	13.2
ظهور السحب	1692	1634	1835	1741	1876	1561	1921	1942	1822	1542	1874

-١٠٦- في روضة للأطفال اختيرت عينة عشوائية من 20 طفل وتم تسجيل درجة كل طفل في الامتحان وكذلك عمره والبيانات معطاة في الجدول التالي :

العمر	6.50	6.75	7.0	7.50	7.50	7.50	7.50	7.75	8.0	8.00
الدرجة	16	28	46	14	41	10	56	43	15	21
العمر	8.25	8.25	8.75	9.00	9.75	9.50	9.50	9.75	9.00	9.00
الدرجة	28	57	36	71	47	66	71	61	60	72

استخدم سبيرمان لاختبار العلاقة بين العمر والدرجة عند مستوى معنوية $\alpha = 0.5$.

REFERENCES المراجع

أولا : المراجع العربية

١-أحمد عبادة سرحان، (١٩٦٨)، مقدمة في طرق التحليل الإحصائي، معهد الدراسات و البحوث الإحصائية-القاهرة .

٢-أحمد عبادة سرحان، (١٩٨٣)، تصميم التجارب و تحليلها، دار الكتب الجامعية-القاهرة-جمهورية مصر العربية .

٣-أحمد فؤاد غالب و آخرون، (١٩٩٣)، الرياضة لدارسى العلوم الحيوية-ترجمة لكتاب جاديش س. أريا و روبين و. لاردنر-الدار الدولية للنشر و التوزيع-القاهرة-جمهورية مصر العربية .

٤-أنيس إسماعيل كنجو، (١٩٩٣)، الإحصاء و الاحتمال-جامعة الملك سعود-عماده شؤون المكتبات .

٥-بدرية شوقى عبد الوهاب و محمد كامل الشربينى، (١٩٨٤)، المبادئ الأولية في الإحصاء -ترجمة لكتاب بول ج. هويل-الطبعة الرابعة-دار جون وايلى وأبنائه .

٦-جلال الصياد، (١٩٨٨)، نظرية الاحتمالات-الطبعة الثانية-دار الشروق-جدة-المملكة العربية السعودية .

٧-جلال مصطفى الصياد و محمد الدسوقى حبيب، (١٩٩٠)، مقدمة في الطرق الإحصائية -الطبعة الثانية-تهامة-جدة-المملكة العربية السعودية .

٨-ذكريا الشربينى، (١٩٩٥)، الإحصاء و تصميم التجارب في البحوث النفسية والتربوية و الاجتماعية-مكتبة الأنجلو المصرية-القاهرة .

٩-ربيع ذكى عامر، (١٩٨٩)، تحليل الانحدار-أساليبه و تطبيقاته العملية باستخدام البرنامج الجاهز SPSS/PC+ -معهد الدراسات و البحوث الإحصائية-جامعة القاهرة .

١٠-سعدية حافظ منتصر، (١٩٨٢)، ملخصات شوم-نظريات و مسائل في الإحصاء و الاقتصاد القياسي-ترجمة لكتاب دومينيك سالفاتور-دار ماكجروهيل-نيويورك .

١١-سمير كامل عاشور و سامية سالم أبو الفتوح، (١٩٩٠)، مقدمة فى الإحصاء التحليلي-معهد الدراسات و البحوث الإحصائية-جامعة القاهرة .

١٢-سمير كامل عاشور و سامية سالم أبو الفتوح، (١٩٩٠)، مقدمة فى الإحصاء الوصفي-معهد الدراسات و البحوث الإحصائية-جامعة القاهرة .

١٣-سمير كامل عاشور و سامية سالم أبو الفتوح، (١٩٩٥)، الاختبارات اللامعلمية-معهد الدراسات و البحوث الإحصائية-جامعة القاهرة .

١٤-عدنان بن ماجد عبد الرحمن برى و محمود محمد إبراهيم هندي و أنور أحمد محمد عبد الله، (١٩٩١)، مبادئ الإحصاء و الاحتمالات-عماده شؤون المكتبات-جامعة الملك سعود-المملكة العربية السعودية .

١٥-عفاف الدش، (١٩٩٤)، الإحصاء التطبيقي للتجاريين-الطبعة الثانية-جامعة حلوان-القاهرة .

-محمد صبحي أبو صالح و عدنان محمد عوض، (١٩٨٣)، مقدمة فى الإحصاء-الطبعة الرابعة-دار جون وايلي و أبنائه-نيويورك .

١٧-محمد محمد الطاهر الإمام، (١٩٩٤)، تصميم و تحليل التجارب-دار المريخ-الرياض-المملكة العربية السعودية .

ثانيا : المراجع الأجنبية

1-Bain, L. J. (1992) Introduction to Probability and Mathematical Statistics, Second Edition, Duxbury Press - An Imprint of Wadsworth Publishing Company Belmont, California.

2-Cangelosi, V. E.; Taylor, P. H. and Rice, P. F. (1979) Basic statistics - A Real World Approach, Second Edition, West Publishing Company, New York.

3-Cochran, W. G. (1963) Sampling Techniques, Second Edition, New York : John Willey & Sons, Inc.

3-Daniel, W. W. (1978) Applied Nonparametric Statistic, Houghton Mifflin Company, London.

4-Devore, J. L. (1995) Probability and Statistics for Engineering and the Sciences, Fourth Edition, Duxbury Press-An International Thomson Publishing Company, London.

5-Draper, N. R. and Smith, H. (1981) Applied Regression Analysis, Second Edition, John Wiley & Sons Inc., U.S.A.

6-Frank, H. and Althoen, S. C. (1997) Statistics- Concepts and Applications, Low Price Edition, Cambridge University Press.

7-Freund, J. E. and Williams, F. J. (1972) Elementary Business Statistics, Englewood Cliffs, N. J. Prentice-Hall, Inc.

8-Hamburg, M. (1979) Basic Statistics : A Modern Approach, Second Edition, Harcourt Brace Jovanovich, Inc., New York.

9-Mendenhall, W. (1975) Introduction to Probability and Statistics, Company,Inc. Belmont,California Fourth Edition, Duxbury Press, A Division of Wadsworth Publishing Company,Inc.Belmont,California.

10-Neter, J.; Wasserman, W. and Whitmere, G. A. (1993) Applied Statistics, Fourth Edition, ALLYN AND BACON, London.

11-Owen, F. and Jones, R. (1994) Statistics, Fourth Edition, Pitman Publishing, London.
12-Schefler, W. (1979) Statistics for the Biological Sciences, Second Edition, Addison-Wesley Publishing Company. Inc. Philippines.
13-Walpole, R. E. (1982) Introduction to Statistics, Macmillan Publishing Co. Inc. New York.
14-Weisberg, S. (1980), Applied Linear Regression, John Wiley & Sons Inc., New York, U.S.A.
15-Winer, B. J. Brown, D. R. and Michels, K. M.(1991) Statistical Experimental Design, Third Edition, McGraw-Hill, Inc., New York.
16-Yamane, T. (1967) Elementary Sampling Theory, Prentice-Hall, Inc., Englewood Cliffs, N. J.
17-Yates, F. (1934) Contingency Tables Involving Small Numbers and the χ^2 Test, J. Roy. Statist. Soc., 1,217-235.

الملاحق

ملحق (١) جدول حساب $\sum_{x=0}^{r} b(x; n, p)$ لمتغير عشوائي يتبع توزيع ذي الحدين .

ملحق (٢) جدول حساب $\sum_{x=0}^{r} p(x; \mu)$ لمتغير عشوائي يتبع توزيع بواسون .

ملحق (٣) جدول المساحات تحت المنحنى الطبيعي القياسي $P(0 < Z < z)$.

ملحق (٤) جدول القيم الحرجة t_α لتوزيع t .

ملحق (٥) جدول القيم الحرجة χ_α^2 لتوزيع χ^2.

ملحق (٦) جدول القيم الحرجة $f_\alpha(\nu_1, \nu_2)$ لتوزيع F عند $(\alpha = 0.05)$.

ملحق (٧) جدول القيم الحرجة $f_\alpha(\nu_1, \nu_2)$ لتوزيع F عند $(\alpha = 0.01)$.

ملحق (٨) جدول القيم الحرجة لاختبار **Cochran** لتجانس التباين .

ملحق (٩) جدول القيم الحرجة $q_\alpha(p, \nu)$ لدانكن .

ملحق (١٠) جدول القيم الحرجة c_α لاختبار الاعتدال .

ملحق (١١) جدول القيم الحرجة $d(n, \alpha'), d(n, \alpha'')$ لاختبار إشارة الرتب .

ملحق (١٢) جدول القيم الحرجة لاختبار **Mann-Whitney-Wilcoxon**

ملحق (١٣) جدول القيم الحرجة لاختبار **Kruskal – Wallis** .

ملحق (١٤) جدول القيم الحرجة r_1 السفلي لاختبار الدورات .

ملحق (١٥) جدول القيم الحرجة r_2 العليا لاختبار الدورات .

ملحق (١٦) جدول القيم الحرجة $r_{s,\alpha}^*$ لاختبار سبيرمان .

ملحق (١)

جدول حساب $\sum_{x=0}^{r} b(x;n,p)$ لمتغير عشوائي يتبع توزيع ذي الحدين

a. n=5

r	0.01	0.05	0.1	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.951	.774	.590	.328	.237	.168	.078	.031	.010	.002	.001	.000	.000	.000	.000
1	.999	.977	.919	.737	.633	.528	.337	.188	.087	.031	.016	.007	.000	.000	.000
2	1.00	.999	.991	.942	.896	.837	.683	.500	.317	.163	.104	.058	.009	.001	.000
3	1.00	1.00	1.00	.993	.984	.969	.913	.812	.663	.472	.367	.263	.081	.023	.001
4	1.00	1.00	1.00	1.00	.999	.998	.990	.969	.922	.832	.763	.672	.410	.226	.049

b. n=10

r	0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.904	.599	.349	.107	.056	.028	.006	.001	.000	.000	.000	.000	.000	.000	.000
1	.996	.914	.736	.376	.244	.149	.046	.011	.002	.000	.000	.000	.000	.000	.000
2	1.00	.988	.930	.678	.526	.383	.167	.055	.012	.002	.000	.000	.000	.000	.000
3	1.00	.999	.987	.879	.776	.650	.382	.172	.055	.011	.004	.001	.000	.000	.000
4	1.00	1.00	.998	.967	.922	.850	.633	.377	.166	.047	.020	.006	.000	.000	.000
5	1.00	1.00	1.00	.994	.980	.953	.834	.623	.367	.150	.078	.033	.002	.000	.000
6	1.00	1.00	1.00	.999	.996	.989	.945	.828	.618	.350	.224	.121	.013	.001	.000
7	1.00	1.00	1.00	1.00	1.00	.998	.988	.945	.833	.617	.474	.322	.070	.012	.000
8	1.00	1.00	1.00	1.00	1.00	1.00	.998	.989	.954	.851	.756	.624	.264	.086	.004
9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.994	.972	.944	.893	.651	.401	.096

تابع ملحق (١)

جدول حساب $\sum_{x=0}^{r} b(x;n,p)$ لمتغير عشوائي يتبع توزيع ذي الحدين

c. n=15

	p														
r	0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.860	.463	.206	.035	.013	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.990	.829	.549	.167	.080	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
2	1.00	.964	.816	.398	.236	.127	.027	.004	.000	.000	.000	.000	.000	.000	.000
3	1.00	.995	.944	.648	.461	.297	.091	.018	.002	.000	.000	.000	.000	.000	.000
4	1.00	.999	.987	.836	.686	.515	.217	.059	.009	.001	.000	.000	.000	.000	.000
5	1.00	1.00	.998	.939	.852	.722	.403	.151	.034	.004	.001	.000	.000	.000	.000
6	1.00	1.00	1.00	.982	.943	.869	.610	.304	.095	.015	.004	.001	.000	.000	.000
7	1.00	1.00	1.00	.996	.983	.950	.787	.500	.213	.050	.017	.004	.000	.000	.000
8	1.00	1.00	1.00	.999	.996	.985	.905	.696	.390	.131	.057	.018	.000	.000	.000
9	1.00	1.00	1.00	1.00	.999	.996	.966	.849	.597	.278	.148	.061	.002	.000	.000
10	1.00	1.00	1.00	1.00	1.00	.999	.991	.941	.783	.485	.314	.164	.013	.001	.000
11	1.00	1.00	1.00	1.00	1.00	1.00	.998	.982	.909	.703	.539	.352	.056	.005	.000
12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.996	.973	.873	.764	.602	.184	.036	.000
13	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.995	.965	.920	.833	.451	.171	.010
14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.995	.987	.965	.794	.537	.140

المصدر : عن [Devore(1995)]

تابع ملحق (١)

جدول حساب $\sum_{x=0}^{r} b(x ; n, p)$ لمتغير عشوائي يتبع توزيع ذي الحدين

d. n=20

r \ p	0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.818	.358	.122	.012	.003	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.983	.736	.392	.069	.024	.008	.001	.000	.000	.000	.000	.000	.000	.000	.000
2	.999	.925	.677	.206	.091	.035	.004	.000	.000	.000	.000	.000	.000	.000	.000
3	1.00	.984	.867	.411	.225	.107	.016	.001	.000	.000	.000	.000	.000	.000	.000
4	1.00	.997	.957	.630	.415	.238	.051	.006	.000	.000	.000	.000	.000	.000	.000
5	1.00	1.00	.989	.804	.617	.416	.126	.021	.002	.000	.000	.000	.000	.000	.000
6	1.00	1.00	.998	.913	.786	.608	.250	.058	.006	.000	.000	.000	.000	.000	.000
7	1.00	1.00	1.00	.968	.898	.772	.416	.132	.021	.001	.000	.000	.000	.000	.000
8	1.00	1.00	1.00	.990	.959	.887	.596	.252	.057	.005	.001	.000	.000	.000	.000
9	1.00	1.00	1.00	.997	.986	.952	.755	.412	.128	.017	.004	.001	.000	.000	.000
10	1.00	1.00	1.00	.999	.996	.983	.872	.588	.245	.048	.014	.003	.000	.000	.000
11	1.00	1.00	1.00	1.00	.999	.995	.943	.748	.404	.113	.041	.010	.000	.000	.000
12	1.00	1.00	1.00	1.00	1.00	.999	.979	.868	.584	.228	.102	.032	.000	.000	.000
13	1.00	1.00	1.00	1.00	1.00	1.00	.994	.942	.750	.392	.214	.087	.002	.000	.000
14	1.00	1.00	1.00	1.00	1.00	1.00	.998	.979	.874	.584	.383	.196	.011	.000	.000
15	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.994	.949	.762	.585	.370	.043	.003	.000
16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.984	.893	.775	.589	.133	.016	.000
17	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.996	.965	.909	.794	.323	.075	.001
18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.992	.976	.931	.608	.264	.017
19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.997	.988	.878	.642	.182

تابع ملحق (١) جدول حساب $\sum_{x=0}^{r} b(x;n,p)$ لمتغير عشوائي يتبع توزيع ذي الحدين

d. n=25

r	p=0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
0	.778	.277	.072	.004	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
1	.974	.642	.271	.027	.007	.002	.000	.000	.000	.000	.000	.000	.000	.000	.000
2	.998	.873	.537	.098	.032	.009	.000	.000	.000	.000	.000	.000	.000	.000	.000
3	1.00	.966	.764	.234	.096	.033	.002	.000	.00	.000	.000	.000	.000	.000	.000
4	1.00	.993	.902	.421	.214	.090	.009	.000	.000	.000	.000	.000	.000	.000	.000
5	1.00	.999	.967	.617	.378	.193	.029	.002	.0	.000	.000	.000	.000	.000	.000
6	1.00	1.00	.991	.780	.561	.341	.074	.007	.000	.000	.000	.000	.000	.000	.000
7	1.00	1.00	.998	.891	.727	.512	.154	.022	.001	.000	.000	.000	.000	.000	.000
8	1.00	1.00	1.00	.953	.851	.677	.274	.054	.004	.000	.000	.000	.000	.000	.000
9	1.00	1.00	1.00	.983	.929	.811	.425	.115	.013	.000	.000	.000	.000	.000	.000
10	1.00	1.00	1.00	.994	.970	.902	.586	.212	.034	.002	.000	.000	.000	.000	.000
11	1.00	1.00	1.00	.998	.980	.956	.732	.345	.078	.006	.001	.000	.000	.000	.000
12	1.00	1.00	1.00	1.00	.997	.983	.846	.500	.154	.017	.003	.000	.000	.000	.000
13	1.00	1.00	1.00	1.00	.999	.994	.922	.655	.268	.044	.020	.002	.000	.000	.000
14	1.00	1.00	1.00	1.00	1.00	.998	.966	.788	.414	.098	.030	.006	.000	.000	.000
15	1.00	1.00	1.00	1.00	1.00	1.00	.987	.885	.575	.189	.071	.017	.000	.000	.000
16	1.00	1.00	1.00	1.00	1.00	1.00	.996	.946	.726	.323	.149	.047	.000	.000	.000
17	1.00	1.00	1.00	1.00	1.00	1.00	.999	.978	.846	.488	.273	.109	.002	.000	.000
18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.993	.926	.659	.439	.220	.009	.000	.000
19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.998	.971	.807	.622	.383	.033	.001	.000
20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.991	.910	.786	.579	.098	.007	.000
21	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.998	.967	.904	.766	.236	.034	.000
22	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.991	.968	.902	.463	.127	.002
23	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.998	.993	.973	.729	.358	.026
24	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.996	.928	.723	.222

ملحق (٢) :

جدول حساب $\sum_{x=0}^{r} p(x;\mu)$ لمتغير عشوائي يتبع توزيع بواسون

		μ									
		.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
r	0	.905	.819	.741	.670	.607	.549	.497	.449	.407	.368
	1	.995	.982	.963	.938	.910	.878	.844	.809	.772	.736
	2	1.00	.999	.996	.992	.986	.977	.966	.953	.937	.920
	3		1.00	1.00	.999	.998	.997	.994	.991	.987	.981
	4				1.00	1.00	1.00	.999	.999	.998	.996
	5							1.00	1.00	1.00	.999
	6										1.00

المصدر عن : [Devore (1995)]

تابع ملحق (٢) : جدول حساب $\sum_{x=0}^{r} p(x;\mu)$ لمتغير عشوائي يتبع توزيع بواسون

	μ										
	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	15.0	20.0
0	.135	.050	.018	.007	.002	.001	.000	.000	.000	.000	.000
1	.406	.199	.092	.040	.017	.007	.003	.001	.000	.000	.000
2	.677	.423	.238	.125	.062	.030	.014	.006	.003	.000	.000
3	.857	.647	.433	.265	.151	.082	.042	.021	.010	.000	.000
4	.947	.815	.629	.440	.285	.173	.100	.055	.029	.001	.000
5	.983	.916	.785	.616	.446	.301	.191	.116	.067	.003	.000
6	.995	.966	.889	.762	.606	.450	.313	.207	.130	.008	.000
7	.999	.988	.949	.867	.744	.599	.453	.324	.220	.018	.001
8	1.00	.996	.979	.932	.847	.729	.593	.456	.333	.037	.002
9		.999	.992	.968	.916	.830	.717	.587	.458	.070	.005
10		1.00	.997	.986	.957	901	.816	.706	.583	.118	.011
11			.999	.995	.980	.947	.888	.803	.697	.185	.021
12			1.00	.998	.991	.973	.936	.876	.792	.268	.039
13				.999	.996	.987	.966	.926	.864	.363	.066
14				1.00	.999	.994	.983	.959	.917	.466	.105
15					.999	.998	.992	.978	.951	.568	.157
16					1.00	.999	.996	.989	.973	.664	.221
17						1.00	.998	.995	.986	.749	.297
18							.999	.998	.993	.819	.381
19							1.00	.999	.997	.875	.470
20								1.00	.998	.917	.559
21									.999	.947	.644
22									1.00	.967	.721
23										.981	.787
24										.989	.843
25										.994	.888
26										.997	.922
27										.998	.948
28										.999	.966
29										1.00	.978
30											.987
31											.992
32											.995
33											.997
34											.999
35											.999
36											1.00

ملحق (٣)

جدول المساحات تحت المنحنى الطبيعي القياسي

$P(0<Z<z)$

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

المصدر : عن **[Daniel (1978)]**

ملحق (٤)

جدول القيم الحرجة t_α لتوزيع **t**

	α						
ν	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.8[illegible]	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

المصدر : عن [Devore (1995)]

ملحق (٥)

جدول القيم الحرجة χ^2_α لتوزيع χ^2

	α									
ν	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.724	26.755
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.735	27.687	29.817
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.600	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.577	32.799
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.407	7.564	8.682	10.085	24.769	27.587	30.190	33.408	35.716
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.843	7.632	8.906	10.117	11.651	27.203	30.143	32.852	36.190	38.580
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.523	13.120	14.611	16.473	34.381	37.652	40.646	44.313	46.925
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.807	12.878	14.573	16.151	18.114	36.741	40.113	43.194	46.962	49.642
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.120	14.256	16.147	17.708	19.768	39.087	42.557	45.772	49.586	52.333
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.457	15.655	17.538	19.280	21.433	41.422	44.985	48.231	52.190	55.000
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.814	17.073	19.046	20.866	23.110	43.745	47.400	50.724	54.774	57.646
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.191	18.508	20.569	22.465	24.796	46.059	49.802	53.203	57.340	60.272
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.584	19.960	22.105	24.075	26.492	48.363	52.192	55.667	59.891	62.880
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.994	21.425	23.654	25.695	28.196	50.660	54.572	58.119	62.426	65.473
40	20.706	22.164	24.433	26.509	29.050	51.805	55.758	59.342	63.691	66.766

المصدر : عن [Devore(1995)]

ملحق (٦)

جدول القيم الحرجة $f_{\alpha}(\nu_1, \nu_2)$ لتوزيع F عند $(\alpha = 0.05)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9..55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.636
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.07
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	·1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.58	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.48	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.84	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

المصدر : عن [Devore (1995)]

ملحق (٧)

جدول القيم الحرجة $f_{\alpha}(\nu_1, \nu_2)$ لتوزيع F عند $(\alpha = 0.01)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.57	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.369
13	9.07	6.70	5.74	5.21	4.86	4.62	4.41	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2..89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.86	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	3.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	3.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	3.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	3.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

ملحق (٨)

جدول القيم الحرجة لاختبار Cochran لتجانس التباين

df for S_j^2	$1-\alpha$	k=number of variances										
		2	3	4	5	6	7	8	9	10	15	20
1	.95	.9985	.9669	.9065	.8412	.7808	.7271	.6798	.6385	.6020	.4709	.3894
	.99	.9999	.9933	.9676	.9279	.8828	.8376	.7945	.7544	.7175	.5747	.4799
2	.95	.9750	.8709	.7679	.6838	.6161	.5612	.5157	.4775	.4450	.3346	.2705
	.99	.9950	.9423	.8643	.7885	.7218	.6644	.6152	.5727	.5358	.4069	.3297
3	.95	.9392	.7977	.6841	.5981	.5321	.4800	.4377	.4027	.3733	.2758	.2205
	.99	.9794	.8831	.7814	.6957	.6258	.5685	.5209	.4810	.4469	.3317	.2654
4	.95	.9057	.7457	.6287	.5441	.4803	.4307	.3910	.3584	.3311	.2419	.1921
	.99	.9586	.8335	.7212	.6329	.5635	.5080	.4627	.4251	.3934	.2882	.2288
5	.95	.8772	.7071	.5895	.5065	.4447	.3974	.3595	.3286	.3029	.2195	.1735
	.99	.9373	.7933	.6761	.5875	.5195	.4659	.4226	.3870	.3572	.2593	.2048
6	.95	.8534	.6771	.5598	.4783	.4184	.3726	.3362	.3067	.2823	.2034	.1602
	.99	.9172	.7606	.6410	.5531	.4866	.4347	.3932	.3592	.3308	.2386	.1877
7	.95	.8332	.6530	.5365	.4564	.3980	.3535	.3185	.2901	.2666	.1911	.1501
	.99	.8988	.7335	.6129	.5259	.4608	.4105	.3704	.3378	.3106	.2228	.1748
8	.95	.8159	.6333	.5175	.4387	.3817	.3384	.3043	.2768	.2541	.1815	.1422
	.99	.8823	.7107	.5897	.5037	.4401	.3911	.3522	.3207	.2945	.2104	.1646
9	.95	.8010	.6167	.5017	.4241	.3682	.3259	.2926	.2659	.2439	.1736	.1357
	.99	.8674	.6912	.5702	.4854	.4229	.3751	.3373	.3067	.2813	.2002	.1567
16	.95	.7341	.5466	.4366	.3645	.3135	.2756	.2462	.2226	.2032	.1429	.1108
	.99	.7949	.6059	.4884	.4094	.3529	.3105	.2779	.2514	.2297	.1612	.1248
36	.95	.6602	.4748	.3720	.3066	.2612	.2278	.2022	.1820	.1655	.1144	.0879
	.99	.7067	5153	.4057	.3351	.2858	.2494	.2214	.1992	.1811	.1251	.0960
144	.95	.5813	.4031	.3093	.2513	.2119	.1833	.1616	.1446	.1308	.0889	.0675
	.99	.6062	.4230	.3251	.2644	.2229	.1929	.1700	.1521	.1376	.0934	.0709

المصدر : عن [Winer et al (1991)]

ملحق (٩)

جدول القيم الحرجة $q_{\alpha}(p,\nu)$ لدانكن

		m									
ν	α	2	3	4	5	6	7	8	9	10	11
5	.05	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17
	.01	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48
6	.05	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65
	.01	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30
7	.05	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30
	.01	4.95	5.92	6.54	7.01	7.37	7.68	6.94	8.17	8.37	8.55
8	.05	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05
	.01	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03
9	.05	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87
	.01	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65
10	.05	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72
	.01	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36
11	.05	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61
	.01	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13
12	.05	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51
	.01	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94
13	.05	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43
	.01	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79
14	.05	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36
	.01	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66
15	.05	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31
	.01	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55
16	.05	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26
	.01	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46
17	.05	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21
	.01	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38
18	.05	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17
	.01	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31
19	.05	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14
	.01	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25
20	.05	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11
	.01	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19
24	.05	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01
	.01	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02
30	.05	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92
	.01	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85
40	.05	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82
	.01	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69
60	.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73
	.01	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53
120	.05	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64
	.01	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37
	.05	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55
	.01	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23

تابع ملحق (٩)

جدول القيم الحرجة $q_{\alpha}(p, \nu)$ لدانكن

m										
12	13	14	15	16	17	18	19	20	α	ν
7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21	.05	5
10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93	.01	
6.79	6.92	7.03	7.14	7.24	7.34	7.43	7.51	7.59	.05	6
9.48	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54	.01	
6.43	6.55	6.66	6.76	6.85	6.94	7.02	7.10	7.17	.05	7
8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65	.01	
6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87	.05	8
8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03	.01	
5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.64	.05	9
7.78	7.91	8.03	8.13	8.23	8.33	8.41	8.49	8.57	.01	
5.83	5.93	6.03	6.11	6.19	6.27	6.34	6.40	6.47	.05	10
7.49	7.60	7.71	7.81	7.91	7.99	8.08	8.15	8.23	.01	
5.71	5.81	5.90	5.98	6.06	6.13	6.20	6.27	6.33	.05	11
7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95	.01	
5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21	.05	12
7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73	.01	
5.53	5.63	5.71	5.79	5.86	5.93	5.99	6.05	6.11	.05	13
6.90	7.01	7.10	7.19	7.27	7.35	7.42	7.48	7.55	.01	
5.46	5.55	5.64	5.71	5.79	5.85	5.91	5.97	6.03	.05	14
6.77	6.87	6.96	7.05	7.13	7.20	7.27	7.33	7.39	.01	
5.40	5.49	5.57	5.65	5.72	5.78	5.85	5.90	5.96	.05	15
6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26	.01	
5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90	.05	16
6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15	.01	
5.31	5.39	5.47	5.54	5.61	5.67	5.73	5.79	5.84	.05	17
6.48	6.57	6.66	6.73	6.81	6.87	6.94	7.00	7.05	.01	
5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79	.05	18
6.41	6.50	6.58	6.65	6.73	6.79	6.85	6.91	6.97	.01	
5.23	5.31	5.39	5.46	5.53	5.59	5.65	5.70	5.75	.05	19
6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89	.01	
5.20	5.28	5.36	5.43	5.49	5.55	5.61	5.66	5.71	.05	20
6.28	6.37	6.45	6.52	6.59	6.65	6.71	6.77	6.82	.01	
5.10	5.18	5.25	5.32	5.38	5.44	5.49	5.55	5.59	.05	24
6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61	.01	
5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.47	.05	30
5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41	.01	
4.90	4.98	5.04	5.11	5.16	5.22	5.27	5.31	5.36	.05	40
5.76	5.83	5.90	5.96	6.02	6.07	6.12	6.16	6.21	.01	
4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24	.05	60
5.60	5.67	5.73	5.78	5.84	5.89	5.93	5.97	6.01	.01	
4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13	.05	120
5.44	5.50	5.56	5.61	5.66	5.71	5.75	5.79	5.83	.01	
4.62	4.68	5.74	4.80	4.85	4.89	4.93	4.97	5.01	.05	∞
5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65	.01	

ملحق (١٠)

جدول القيم الحرجة c_α لاختبار الإعتدال

	α		
	.1	.05	.01
5	.9033	.8804	.8320
10	.9347	.9180	.8804
15	.9506	.9383	.9110
20	.9600	.9503	.9290
n 25	.9662	.9582	.9408
30	.9707	.9639	.9490
40	.9767	.9715	.9597
50	.9807	.9764	.9664
60	.9835	.9799	.9710
75	.9865	.9835	.9757

المصدر عن : **[Daniel (1978)]**

ملحق (١١)

جدول القيم الحرجة $d(n, \alpha''), d(n, \alpha')$ لاختبار إشارة الرتب

n	d	Confidence coefficient	α''	α'
3	1	.750	.250	.125
4	1	.875	.125	.063
5	1	.938	.062	.031
	2	.875	.125	.063
6	1	.969	.031	.016
	2	.937	.063	.031
	3	.906	.094	.047
	4	.844	.156	.078
7	1	.984	.016	.008
	2	.969	.031	.016
	4	.922	.078	.039
	5	.891	.109	.055
8	1	.992	.008	.004
	2	.984	.016	.008
	4	.961	.039	.020
	5	.945	.055	.027
	6	.922	.078	.039
	7	.891	.109	.055
9	2	.992	.008	.004
	3	.988	.012	.006
	6	.961	.039	.020
	7	.945	.055	.027
	9	.902	.098	.049
	10	.871	.129	.065
10	4	.990	.010	.005
	5	.986	.014	.007
	9	.951	.049	.024
	10	.936	.064	.032
	11	.916	.084	.042
	12	.895	.105	.053
11	6	.990	.010	.005
	7	.986	.014	.007
	11	.958	.042	.021
	12	.946	.054	.027
	14	.917	.083	.042
	15	.898	.102	.051
12	8	.991	.009	.005
	9	.988	.012	.006
	14	.958	.042	.021
	15	.948	.052	.026
	18	.908	.092	.046
	19	.890	.110	.055
13	10	.992	.008	.004
	11	.990	.010	.005
	18	.952	.048	.024
	19	.943	.057	.029
	22	.906	.094	.047
	23	.890	.110	.055

المصدر : عن [Daniel (1978)]

تابع : ملحق (١١)

جدول القيم الحرجة $d(n, \alpha'), d(n, \alpha'')$ لاختبار إشارة الرتب

n	d	Confidence coefficient	α''	α'
14	13	.991	.009	.004
	14	.989	.011	.005
	22	.951	.049	.025
	23	.942	.058	.029
	26	.909	.091	.045
	27	.896	.104	.052
15	16	.992	.008	.004
	17	.990	.010	.005
	26	.952	.048	.024
	27	.945	.055	.028
	31	.905	.095	.047
	32	.893	.107	.054
16	20	.991	.009	.005
	21	.989	.011	.006
	30	.956	.044	.022
	31	.949	.051	.025
	36	.907	.093	.047
	37	.895	.105	.052
17	24	.991	.009	.005
	25	.989	.011	.006
	35	.955	.045	.022
	36	.949	.051	.025
	42	.902	.098	.049
	43	.891	.109	.054
18	28	.991	.009	.005
	29	.990	.010	.005
	41	.952	.048	.024
	42	.946	.054	.027
	48	.901	.099	.049
	49	.892	.108	.054
19	33	.991	.009	.005
	34	.989	.011	.005
	47	.951	.049	.025
	48	.945	.055	.027
	54	.904	.096	.048
	55	.896	.104	.052
20	38	.991	.009	.005
	39	.989	.011	.005
	53	.952	.048	.024
	54	.947	.053	.027
	61	.903	.097	.049
	62	.895	.105	.053
21	43	.991	.009	.005
	44	.990	.010	.005
	59	.954	.046	.023
	60	.950	.050	.025
	68	.904	.096	.048
	69	.897	.103	.052

المصدر : عن [Daniel (1978)]

تابع : ملحق (١١)

جدول القيم الحرجة $d(n,\alpha'), d(n,\alpha'')$ لاختبار الإشارة

n	d	Confidence coefficient	α''	α'
22	49	.991	.009	.005
	50	.990	.010	.005
	66	.954	.046	.023
	67	.950	.050	.025
	76	.902	.098	.049
	77	.895	.105	.053
23	55	.991	.009	.005
	56	.990	.010	.005
	74	.952	.048	.024
	75	.948	.052	.026
	84	.902	.098	.049
	85	.895	.105	.052
24	62	.990	.010	.005
	63	.989	.011	.005
	82	.951	.049	.025
	83	.947	.053	.026
	92	.905	.095	.048
	93	.899	.101	.051
25	69	.990	.010	.005
	70	.989	.011	.005
	90	.952	.048	.024
	91	.948	.052	.026
	101	.904	.096	.048
	102	.899	.101	.051

المصدر : عن **[Daniel (1978)]**

ملحق (١٢)

جدول القيم الحرجة لاختبار

Mann-Whitney-Wilcoxon

n_1	α	n_2=2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	.005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
2	.01	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2	2
	.025	0	0	0	0	0	0	1	1	1	1	2	2	2	2	2	3	3	3	3
	.05	0	0	0	1	1	1	2	2	2	2	3	3	4	4	4	4	5	5	5
	.10	0	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8
	001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	.005	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	3	3	4	4
3	.01	0	0	0	0	0	1	1	2	2	2	3	3	3	4	4	5	5	5	6
	.025	0	0	0	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9
	.05	0	1	1	2	3	3	4	5	5	6	6	7	8	8	9	10	10	11	12
	.10	1	2	2	3	4	5	6	6	7	8	9	10	11	11	12	13	14	15	16
	001	0	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4
	.005	0	0	0	0	1	1	2	2	3	3	4	4	5	6	6	7	7	8	9
4	.01	0	0	0	1	2	2	3	4	4	5	6	6	7	9	8	9	10	10	11
	.025	0	0	1	2	3	4	5	5	6	7	8	9	10	11	12	12	13	14	15
	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19
	.10	1	2	4	5	6	7	8	10	11	12	13	14	16	17	18	19	21	22	23
	001	0	0	0	0	0	0	1	2	2	3	3	4	4	5	6	6	7	8	8
	.005	0	0	0	1	2	2	3	4	5	6	7	8	8	9	10	11	12	13	14
5	.01	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	.025	0	1	2	3	4	6	7	8	9	10	12	13	14	15	16	18	19	20	21
	.05	1	2	3	5	6	7	9	10	12	13	14	16	17	19	20	21	23	24	26
	.10	2	3	5	6	8	9	11	13	14	16	18	19	21	23	24	26	28	29	31
	001	0	0	0	0	0	0	2	3	4	5	5	6	7	8	9	10	11	12	13
	.005	0	0	1	2	3	4	5	6	7	8	10	11	12	13	14	16	17	18	19
6	.01	0	0	2	3	4	5	7	8	9	10	12	13	14	16	17	19	20	21	23
	.025	0	2	3	4	6	7	9	11	12	14	15	17	18	20	22	23	25	26	28
	.05	1	3	4	6	8	9	11	13	15	17	18	20	22	24	26	27	29	31	33
	.10	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	35	37	39
	001	0	0	0	0	1	2	3	4	6	7	8	9	10	11	12	14	15	16	17
	.005	0	0	1	2	4	5	7	8	10	11	13	14	16	17	19	20	22	23	25
7	.01	0	1	2	4	5	7	8	10	12	13	15	17	18	20	22	24	25	27	29
	.025	0	2	4	6	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
	.05	1	3	5	7	9	12	14	16	18	20	22	25	27	29	31	34	36	38	40
	.10	2	5	7	9	12	14	17	19	22	24	27	29	32	34	37	39	42	44	47
	001	0	0	0	1	2	3	5	6	7	9	10	12	13	15	16	18	19	21	22
	.005	0	0	2	3	5	7	8	10	12	14	16	18	19	21	23	25	27	29	31
8	.01	0	1	3	5	7	8	10	12	14	16	18	21	23	25	27	29	31	33	35
	.025	1	3	5	7	9	11	14	16	18	20	23	25	27	30	32	35	37	39	42
	.05	2	4	6	9	11	14	16	19	21	24	27	29	32	34	37	40	42	45	48
	.10	3	6	8	11	14	17	20	23	25	28	31	34	37	40	43	46	49	52	55

المصدر : عن [Daniel (1978)]

تابع ملحق (١٢)

جدول القيم الحرجة لاختبار
Mann-Whitney-Wilcoxon

n_1	α	n_2=2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
9	.001	0	0	0	2	3	4	6	8	9	11	13	15	16	18	20	22	24	26	27
	.005	0	1	2	4	6	8	10	12	14	17	19	21	23	25	28	30	32	34	37
	.01	0	2	4	6	8	10	12	15	17	19	22	24	27	29	32	34	37	39	41
	.025	1	3	5	8	11	13	16	18	21	24	27	29	32	35	38	40	43	46	49
	.05	2	5	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55
	.10	3	6	10	13	16	19	23	26	29	32	36	39	42	46	49	53	56	59	63
10	.001	0	0	1	2	4	6	7	9	11	13	15	18	20	22	24	26	28	30	33
	.005	0	1	3	5	7	10	12	14	17	19	22	25	27	30	32	35	38	40	43
	.01	0	2	4	7	9	12	14	17	20	23	25	28	31	34	37	39	42	45	48
	.025	1	4	6	9	12	15	18	21	24	27	30	34	37	40	43	46	49	53	56
	.05	2	5	8	12	15	18	21	25	28	32	35	38	42	45	49	52	56	59	63
	.10	4	7	11	14	18	22	25	29	33	37	40	44	48	52	55	59	63	67	71
11	.001	0	0	1	3	5	7	9	11	13	16	18	21	23	25	28	30	33	35	38
	.005	0	1	3	6	8	11	14	17	19	22	25	28	31	34	37	40	43	46	49
	.01	0	2	5	8	10	13	16	19	23	26	29	32	35	38	42	45	48	51	54
	.025	1	4	7	10	14	17	20	24	27	31	34	38	41	45	48	52	56	59	63
	.05	2	6	9	13	17	20	24	28	32	35	39	43	47	51	55	58	62	66	70
	.10	4	8	12	16	20	24	28	32	37	41	45	49	53	58	62	66	70	74	79
12	.001	0	0	1	3	5	8	10	13	15	18	21	24	26	29	32	35	38	41	43
	.005	0	2	4	7	10	13	16	19	22	25	28	32	35	38	42	45	48	52	55
	.01	0	3	6	9	12	15	18	22	25	29	32	36	39	43	47	50	54	57	61
	.025	2	5	8	12	15	19	23	27	30	34	38	42	46	50	54	58	62	66	70
	.05	3	6	10	14	18	22	27	31	35	39	43	48	52	56	61	65	69	73	78
	.10	5	9	13	18	22	27	31	36	40	45	50	54	59	64	68	73	78	82	87
13	.001	0	0	2	4	6	9	12	15	18	21	24	27	30	33	36	39	43	46	49
	.005	0	2	4	8	11	14	18	21	25	28	32	35	39	43	46	50	54	58	61
	.01	1	3	6	10	13	17	21	24	28	32	36	40	44	48	52	56	60	64	68
	.025	2	5	9	13	17	21	25	29	34	38	42	46	51	55	60	64	68	73	77
	.05	3	7	11	16	20	25	29	34	38	43	48	52	57	62	66	71	76	81	85
	.10	5	10	14	19	24	29	34	39	44	49	54	59	64	69	75	80	85	90	95
14	.001	0	0	2	4	7	10	13	16	20	23	26	30	33	37	40	44	47	51	55
	.005	0	2	5	8	12	16	19	23	27	31	35	39	43	47	51	55	59	64	68
	.01	1	3	7	11	14	18	23	27	31	35	39	44	48	52	57	61	66	70	74
	.025	2	6	10	14	18	23	27	32	37	41	46	51	56	60	65	70	75	79	84
	.05	4	8	12	17	22	27	32	37	42	47	52	57	62	67	72	78	83	88	93
	.10	5	11	16	21	26	32	37	42	48	53	59	64	70	75	81	86	92	98	103
15	.001	0	0	2	5	8	11	15	18	22	25	29	33	37	41	44	48	52	56	60
	.005	0	3	6	9	13	17	21	25	30	34	38	43	47	52	56	61	65	70	74
	.01	1	4	8	12	16	20	25	29	34	38	43	48	52	57	62	67	71	76	81
	.025	2	6	11	15	20	25	30	35	40	45	50	55	60	65	71	76	81	86	91
	.05	4	8	13	19	24	29	34	40	46	51	56	62	67	73	78	84	89	95	101
	.10	6	11	17	23	28	34	40	46	52	58	64	69	75	81	87	93	99	105	111
16	.001	0	0	3	6	9	12	16	20	24	28	32	36	40	44	49	53	57	61	66
	.005	0	3	6	10	14	19	23	28	32	37	42	46	51	56	61	66	71	75	80
	.01	1	4	8	13	17	22	27	32	37	42	47	52	57	62	67	72	77	83	88
	.025	2	7	12	16	22	27	32	38	43	48	54	60	65	71	76	82	87	93	99
	.05	4	9	15	20	26	31	37	43	49	55	61	66	72	78	84	90	96	102	108
	.10	6	12	18	24	30	37	43	49	55	62	68	75	81	87	94	100	107	113	120

تابع ملحق (١٢)

جدول القيم الحرجة لاختبار

Mann-Whitney-Wilcoxon

n_1	P	$n_2=2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	1	3	6	10	14	18	22	26	30	35	39	44	48	53	58	62	67	71
	.005	0	3	7	11	16	20	25	30	35	40	45	50	55	61	66	71	76	82	87
17	.01	1	5	9	14	19	24	29	34	39	45	50	56	61	67	72	78	83	89	94
	.025	3	7	12	18	23	29	35	40	46	52	58	64	70	76	82	88	94	100	106
	.05	4	10	16	21	27	34	40	46	52	58	65	71	78	84	90	97	103	110	116
	.10	7	13	19	26	32	39	46	53	59	66	73	80	86	93	100	107	114	121	128
	.001	0	1	4	7	11	15	19	24	28	33	38	43	47	52	57	62	67	72	77
	.005	0	3	7	12	17	22	27	32	38	43	48	54	59	65	71	76	82	88	93
18	.01	1	5	10	15	20	25	31	37	42	48	54	60	66	71	77	83	89	95	101
	.025	3	8	13	19	25	31	37	43	49	56	62	68	75	81	87	94	100	107	113
	.05	5	10	17	23	29	36	42	49	56	62	69	76	83	89	96	103	110	117	124
	.10	7	14	21	28	35	42	49	56	63	70	78	85	92	99	107	114	121	129	136
	.001	0	1	4	8	12	16	21	26	30	35	41	46	51	56	61	67	72	78	83
	.005	1	4	8	13	18	23	29	34	40	46	52	58	64	70	75	82	88	94	100
19	.01	2	5	10	16	21	27	33	39	45	51	57	64	70	76	83	89	95	102	108
	.025	3	8	14	20	26	33	39	46	53	59	66	73	79	86	93	100	107	114	120
	.05	5	11	18	24	31	38	45	52	59	66	73	81	88	95	102	110	117	124	131
	.10	8	15	22	29	37	44	52	59	67	74	82	90	98	105	113	121	129	136	144
	.001	0	1	4	8	13	17	22	27	33	38	43	49	55	60	66	71	77	83	89
	.005	1	4	9	14	19	25	31	37	43	49	55	61	68	74	80	87	93	100	106
20	.01	2	6	11	17	23	29	35	41	48	54	61	68	74	81	88	94	101	108	115
	.025	3	9	15	21	28	35	42	49	56	63	70	77	84	91	99	106	113	120	128
	.05	5	12	19	26	33	40	48	55	63	70	78	85	93	101	108	116	124	131	139
	.10	8	16	23	31	39	47	55	63	71	79	87	95	103	111	120	128	136	144	152

ملحق (١٣)

جدول القيم الحرجة لاختبار Kruskal – Wallis

Sample Sizes				
n_1	n_2	n_3	Critical value	α
2	1	1	2.7000	0.500
2	2	1	3.6000	0.200
2	2	2	4.5714	0.067
			3.7143	0.200
3	1	1	3.2000	0.300
3	2	1	4.2857	0.100
			3.8571	0.133
3	2	2	5.3572	0.029
			4.7143	0.048
			4.5000	0.067
			4.4643	0.105
3	3	1	5.1429	0.043
			4.5714	0.100
			4.000	0.129
3	3	2	6.2500	0.011
			5.3611	0.032
			5.1389	0.061
			4.5556	0.100
			4.2500	0.121
3	3	3	7.2000	0.004
			6.4889	0.011
			5.6889	0.029
			5.6000	0.050
			5.0667	0.086
			4.6222	0.100
4	1	1	3.5714	0.200
4	2	1	4.8214	0.057
			4.5000	0.076
			4.0179	0.114
4	2	2	6.0000	0.014
			5.3333	0.033
			5.1250	0.052
			4.4583	0.100
			4.1667	0.105
4	3	1	5.8333	0.021
			5.2083	0.050
			5.0000	0.057
			4.0556	0.093
			4.8889	0.129
4	3	2	6.4444	0.008
			6.3000	0.011
			5.4444	0.046
			5.4000	0.051
			4.5111	0.098
			4.4444	0.102
4	3	3	6.7455	0.010
			6.7091	0.013
			5.7909	0.046
			5.7273	0.050
			4.7091	0.092

Sample Sizes				
n_1	n_2	n_3	Critical value	α
			4.7000	0.101
4	4	1	6.6667	0.010
			6.1667	0.022
			4.9667	0.048
			4.8667	0.054
			4.1667	0.082
			4.0667	0.102
4	4	2	7.0364	0.006
			6.8727	0.011
			5.4545	0.046
			5.2364	0.052
			4.5545	0.098
			4.4455	0.103
4	4	3	7.1439	0.010
			7.1364	0.011
			5.5985	0.049
			5.5758	0.051
			4.5455	0.099
			4.4773	0.102
4	4	4	7.6538	0.008
			7.5385	0.011
			5.6923	0.049
			5.6538	0.054
			4.6539	0.097
			4.5001	0.104
5	1	1	3.8571	0.143
5	2	1	5.2500	0.036
			5.0000	0.048
			4.4500	0.071
			4.2000	0.095
			4.0500	0.119
5	2	2	6.5333	0.008
			6.1333	0.013
			5.1600	0.034
			5.0400	0.056
			4.3733	0.090
			4.2933	0.122
5	3	1	6.4000	0.012
			4.9600	0.048
			4.8711	0.052
			4.0178	0.095
			3.8400	0.123
5	3	2	6.9091	0.009
			6.8218	0.010
			5.2509	0.049
			5.1055	0.052
			4.6509	0.091
			4.4945	0.101
5	3	3	7.0788	0.009
			6.9818	0.011

تابع : ملحق (١٣)

جدول القيم الحرجة لاختبار Kruskal – Wallis

Sample Sizes					Sample Sizes				
n_1	n_2	n_3	Critical value	α	n_1	n_2	n_3	Critical value	α
5	3	3	5.6485	0.049	5	5	1	6.8364	0.011
			5.5152	0.051				5.1273	0.046
			4.5333	0.097				4.9091	0.053
			4.4121	0.109				4.1091	0.086
5	4	1	6.9545	0.008				4.0364	0.105
			6.8400	0.011	5	5	2	7.3385	0.010
			4.9855	0.044				7.2692	0.010
			4.8600	0.056				5. 3385	0.047
			3.9873	0.098				5.2462	0.051
			3.9600	0.102				4.6231	0.097
5	4	2	7.2045	0.009				4.5077	0.100
			7.1182	0.010	5	5	3	7.5780	0.010
			5.2727	0.049				7.5429	0.010
			5.2682	0.050				5.7055	0.046
			4.5409	0.098				5.6264	0.051
			4.5182	0.101				4.5451	0.100
5	4	3	7.4449	0.010				4.5363	0.102
			7.3949	0.011	5	5	4	7.8229	0.010
			5.6564	0.049				7.7914	0.010
			5.6308	0.050				5.6657	0.049
			4.5487	0.099				5.6429	0.050
			4.5231	0.103				4.5229	0.099
5	4	4	7.7604	0.009				4.5200	0.101
			7.7440	0.011	5	5	5	8.000	0.009
			5.6571	0.049				7.9800	0.010
			5.6176	0.050				5.7800	0.049
			4.6187	0.100				5.6600	0.051
			4.5527	0.102				4.5600	0.100
5	5	1	7.3091	0.009				4.5000	0.102

المصدر : عن[Daniel (1978)]

ملحق (١٤)

جدول القيم الحرجة r_1 السفلي لاختبار الدورات

n_1 \ n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

المصدر : عن [Daniel (1978)]

ملحق (١٥)

جدول القيم الحرجة r_2 العليا لاختبار الدورات

n_1 \ n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13						15	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

المصدر : عن [Daniel (1978)]

ملحق (١٦)

جدول القيم الحرجة $r^*_{s,\alpha}$ لاختبار سبيرمان

n	.001	.005	.010	.025	.050	.100
4	--	--	--	--	.8000	.8000
5	--	--	.9000	.9000	.8000	.7000
6	--	.9429	.8857	.8286	.7714	.6000
7	.9643	.8929	.8571	.7450	.6786	.5357
8	.9286	.8571	.8095	.7143	.6190	.5000
9	.9000	.8167	.7667	.6833	.5833	.4667
10	.8667	.7818	.7333	.6364	.5515	.4424
11	.8364	.7545	.7000	.6091	.5273	.4182
12	.8182	.7273	.6713	.5804	.4965	.3986
13	.7912	.6978	.6429	.5549	.4780	.3791
14	.7670	.6747	.6220	.5341	.4593	.3626
15	.7464	.6536	.6000	.5179	.4429	.3500
16	.7265	.6324	.5824	.5000	.2465	.3382
17	.7083	.6152	.5637	.4853	.4118	.3260
18	.6904	.5975	.5480	.4716	.3994	.3148
19	.6737	.5825	.5333	.4579	.3895	.3070
20	.6586	.5684	.5203	.4451	.3789	.2977
21	.6455	.5545	.5078	.4351	.3688	.2909
22	.6318	.5426	.4963	.4241	.3597	.2829
23	.6186	.5306	.4852	.4150	.3518	.2767
24	.6070	.5200	.4748	.4061	.3435	.2704
25	.5962	.5100	.4654	.3977	.3362	.2646
26	.5856	.5002	.4564	.3894	.3299	.2588
27	.5757	.4915	.4481	.3822	.3236	.2540
28	.5660	.4828	.4401	.3749	.3175	.2490
29	.5567	.4744	.4320	.3685	.3113	.2443
30	.5479	.4665	.4251	.3620	.3059	.2400

المصدر : عن [Daniel (1978)]

